

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 13:23:59 : Search time 393.89 Seconds
(without alignments)
3013.345 Million cell updates/sec

Title: US-09-138-735-1_COPY_1232_1825

Perfect score: 594
Sequence: 1 caggtacagcgtacgctt.....aaacgagctccgtgagcata 594

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1746710 seqs, 999096356 residues

Word size : 0
Total number of hits satisfying chosen parameters: 3493420

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Pending_Patents_MN_New.*

1: /cgn1_7/ptodata/1/pna/PCR_NEW_COMB.seq.*
2: /cgn1_7/ptodata/1/pna/US06_NEW_COMB.seq.*
3: /cgn1_7/ptodata/1/pna/US07_NEW_COMB.seq.*
4: /cgn1_7/ptodata/1/pna/US08_NEW_COMB.seq.*
5: /cgn1_7/ptodata/1/pna/US09_NEW_COMB.seq.*
6: /cgn1_7/ptodata/1/pna/US09_NEW_COMB.seq1.*
7: /cgn1_7/ptodata/1/pna/US09_NEW_COMB.seq2.*
8: /cgn1_7/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 594 | 100.0 | 3402 | 6 | US-09-138-735-1 |
| 2 | 22 | 3.7 | 22 | 6 | US-09-138-735-9 |
| 3 | 20 | 3.4 | 91470 | 6 | US-09-803-736-19 |
| 4 | 18 | 3.0 | 18 | 6 | US-09-138-735-5 |
| 5 | 18 | 3.0 | 18 | 6 | US-09-138-735-8 |
| 6 | 18 | 3.0 | 261 | 6 | US-09-866-555-4980 |
| 7 | 18 | 3.0 | 297 | 7 | US-09-863-524-3874 |
| 8 | 18 | 3.0 | 328 | 6 | US-09-764-891-1027 |
| 9 | 18 | 3.0 | 348 | 7 | US-09-540-213-5492 |
| 10 | 18 | 3.0 | 371 | 6 | US-09-667-188A-4578 |
| 11 | 18 | 3.0 | 468 | 7 | US-09-864-761-5238 |
| 12 | 18 | 3.0 | 476 | 6 | US-09-904-809-19378 |
| 13 | 18 | 3.0 | 476 | 6 | US-09-909-627-11623 |
| 14 | 18 | 3.0 | 506 | 7 | US-09-864-761-16893 |
| 15 | 18 | 3.0 | 506 | 7 | US-09-840-145-13709 |
| 16 | 18 | 3.0 | 1431 | 8 | US-60-312-544-4166 |
| 17 | 18 | 3.0 | 1485 | 1 | PCR-US01-16450-789 |
| 18 | 18 | 3.0 | 2817 | 6 | US-09-836-472-1298 |
| 19 | 18 | 3.0 | 15270 | 7 | US-09-764-905-28829 |
| 20 | 18 | 3.0 | 15270 | 7 | US-09-764-905-40566 |
| 21 | 18 | 3.0 | 96424 | 6 | US-09-803-736-1397 |
| 22 | 17 | 2.9 | 241 | 7 | US-09-724-671-11298 |
| 23 | 17 | 2.9 | 241 | 7 | US-09-540-213-52685 |
| 24 | 17 | 2.9 | 264 | 7 | US-09-540-213-44587 |
| 25 | 17 | 2.9 | 300 | 7 | US-09-540-213-33702 |

| | | | | | |
|----|----|-----|--------|----|----------------------|
| 26 | 17 | 2.9 | 301 | 7 | US-09-724-671-8427 |
| 27 | 17 | 2.9 | 303 | 7 | US-09-724-671-8448 |
| 28 | 17 | 2.9 | 322 | 7 | US-09-863-524-3604 |
| 29 | 17 | 2.9 | 353 | 7 | US-09-724-671-5748 |
| 30 | 17 | 2.9 | 396 | 6 | US-09-724-671-4945 |
| 31 | 17 | 2.9 | 398 | 6 | US-09-804-730-15348 |
| 32 | 17 | 2.9 | 402 | 6 | US-09-904-703-4316 |
| 33 | 17 | 2.9 | 404 | 6 | US-09-866-555-6918 |
| 34 | 17 | 2.9 | 419 | 7 | US-09-692-257A-6664 |
| 35 | 17 | 2.9 | 426 | 6 | US-09-804-730-8833 |
| 36 | 17 | 2.9 | 430 | 6 | US-09-904-809-4385 |
| 37 | 17 | 2.9 | 439 | 7 | US-09-881-797-3357 |
| 38 | 17 | 2.9 | 444 | 7 | US-09-724-671-4408 |
| 39 | 17 | 2.9 | 468 | 6 | US-09-758-463-219 |
| 40 | 17 | 2.9 | 474 | 7 | US-09-881-797-3356 |
| 41 | 17 | 2.9 | 481 | 8 | US-60-253-652-24408 |
| 42 | 17 | 2.9 | 492 | 6 | US-09-823-241-6683 |
| 43 | 17 | 2.9 | 500 | 7 | US-09-850-147-15830 |
| 44 | 17 | 2.9 | 501 | 7 | US-09-834-366-4311 |
| 45 | 17 | 2.9 | 514 | 6 | US-09-758-463-239 |
| 46 | 17 | 2.9 | 562 | 7 | US-09-821-837-4107 |
| 47 | 17 | 2.9 | 643 | 7 | US-09-823-327-3606 |
| 48 | 17 | 2.9 | 697 | 6 | US-09-623-746A-60 |
| 49 | 17 | 2.9 | 771 | 1 | PCR-US01-08631-9801 |
| 50 | 17 | 2.9 | 781 | 1 | PCR-US01-08631-9802 |
| 51 | 17 | 2.9 | 790 | 6 | US-09-883-152-20 |
| 52 | 17 | 2.9 | 812 | 8 | US-60-278-561-2143 |
| 53 | 17 | 2.9 | 954 | 1 | PCR-US01-18569-775 |
| 54 | 17 | 2.9 | 1269 | 1 | PCR-US01-08631-10338 |
| 55 | 17 | 2.9 | 1513 | 6 | US-09-898-888-7596 |
| 56 | 17 | 2.9 | 1707 | 1 | PCR-US01-08631-29984 |
| 57 | 17 | 2.9 | 1959 | 8 | US-60-278-561-12846 |
| 58 | 17 | 2.9 | 1970 | 1 | PCR-US01-08631-21764 |
| 59 | 17 | 2.9 | 2205 | 1 | PCR-US01-08631-8851 |
| 60 | 17 | 2.9 | 2211 | 7 | US-09-764-874-12555 |
| 61 | 17 | 2.9 | 2914 | 8 | US-60-278-256-10032 |
| 62 | 17 | 2.9 | 3465 | 8 | US-60-278-233-8890 |
| 63 | 17 | 2.9 | 3617 | 7 | US-09-764-874-12556 |
| 64 | 17 | 2.9 | 3618 | 7 | US-09-764-874-12554 |
| 65 | 17 | 2.9 | 3846 | 7 | US-09-902-540-2520 |
| 66 | 17 | 2.9 | 4815 | 1 | PCR-US01-08631-24178 |
| 67 | 17 | 2.9 | 4815 | 1 | PCR-US01-08631-26102 |
| 68 | 17 | 2.9 | 4826 | 8 | US-60-260-483-126 |
| 69 | 17 | 2.9 | 4835 | 8 | US-60-278-258-4418 |
| 70 | 17 | 2.9 | 5491 | 17 | PCR-US01-08631-21772 |
| 71 | 17 | 2.9 | 6238 | 1 | PCR-US01-08631-4045 |
| 72 | 17 | 2.9 | 6266 | 8 | US-60-278-232-9933 |
| 73 | 17 | 2.9 | 7714 | 6 | US-09-760-457-438 |
| 74 | 17 | 2.9 | 7714 | 6 | US-09-760-457-463 |
| 75 | 17 | 2.9 | 17503 | 7 | US-09-902-540-1114 |
| 76 | 17 | 2.9 | 32189 | 6 | US-09-764-874-3391 |
| 77 | 17 | 2.9 | 32189 | 7 | US-09-764-874-8672 |
| 78 | 17 | 2.9 | 32189 | 7 | US-09-764-874-9623 |
| 79 | 17 | 2.9 | 32189 | 7 | US-09-764-874-10121 |
| 80 | 17 | 2.9 | 46626 | 6 | US-09-803-736-184 |
| 81 | 17 | 2.9 | 80099 | 6 | US-09-803-736-1287 |
| 82 | 17 | 2.9 | 94029 | 6 | US-09-803-736-147 |
| 83 | 17 | 2.9 | 240825 | 6 | US-09-790-288-1 |
| 84 | 16 | 2.7 | 73 | 8 | US-60-253-457-27313 |
| 85 | 16 | 2.7 | 139 | 7 | US-09-864-761-27193 |
| 86 | 16 | 2.7 | 235 | 7 | US-09-724-671-4688 |
| 87 | 16 | 2.7 | 236 | 6 | US-09-863-524-4502 |
| 88 | 16 | 2.7 | 236 | 8 | US-60-253-376-35947 |
| 89 | 16 | 2.7 | 239 | 8 | US-60-253-376-35896 |
| 90 | 16 | 2.7 | 241 | 7 | US-09-853-366-7523 |
| 91 | 16 | 2.7 | 245 | 7 | US-09-877-486-3351 |
| 92 | 16 | 2.7 | 275 | 7 | US-09-912-292-55592 |
| 93 | 16 | 2.7 | 279 | 7 | US-09-853-366-1403 |
| 94 | 16 | 2.7 | 285 | 7 | US-09-859-490-835 |
| 95 | 16 | 2.7 | 290 | 6 | US-09-617-681A-3535 |
| 96 | 16 | 2.7 | 293 | 8 | US-60-252-833-38373 |
| 97 | 16 | 2.7 | 297 | 7 | US-09-867-716-7328 |
| 98 | 16 | 2.7 | 302 | 6 | US-09-783-587-1286 |

| |
|--------------------|
| Sequence 8427, Ap |
| Sequence 8448, Ap |
| Sequence 3604, Ap |
| Sequence 5748, Ap |
| Sequence 4945, Ap |
| Sequence 15348, A |
| Sequence 4316, Ap |
| Sequence 6918, Ap |
| Sequence 6664, Ap |
| Sequence 8033, Ap |
| Sequence 4385, Ap |
| Sequence 3357, Ap |
| Sequence 4408, Ap |
| Sequence 219, App |
| Sequence 3356, Ap |
| Sequence 24408, A |
| Sequence 6683, Ap |
| Sequence 15830, A |
| Sequence 4311, Ap |
| Sequence 239, App |
| Sequence 4107, Ap |
| Sequence 3606, Ap |
| Sequence 60, App1 |
| Sequence 9801, Ap |
| Sequence 9802, Ap |
| Sequence 20, App1 |
| Sequence 2143, Ap |
| Sequence 775, App |
| Sequence 1038, A |
| Sequence 7596, Ap |
| Sequence 12884, A |
| Sequence 12986, A |
| Sequence 21764, A |
| Sequence 8851, Ap |
| Sequence 12555, A |
| Sequence 10032, A |
| Sequence 8890, Ap |
| Sequence 12556, A |
| Sequence 12554, A |
| Sequence 2520, Ap |
| Sequence 24178, A |
| Sequence 26102, A |
| Sequence 126, App |
| Sequence 4418, Ap |
| Sequence 21772, A |
| Sequence 4045, Ap |
| Sequence 9933, Ap |
| Sequence 458, App |
| Sequence 463, App |
| Sequence 1114, Ap |
| Sequence 3291, Ap |
| Sequence 8672, Ap |
| Sequence 9823, Ap |
| Sequence 10121, A |
| Sequence 164, App |
| Sequence 1287, App |
| Sequence 147, App |
| Sequence 1, App1 |
| Sequence 27313, A |
| Sequence 27193, A |
| Sequence 4868, Ap |
| Sequence 4502, Ap |
| Sequence 35947, A |
| Sequence 35986, A |
| Sequence 7623, Ap |
| Sequence 3251, Ap |
| Sequence 52592, A |
| Sequence 1403, Ap |
| Sequence 8355, App |
| Sequence 38373, A |
| Sequence 7328, Ap |
| Sequence 1286, Ap |

| | | | | | | | | | | | | | |
|-------|----|-----|------|---|-----------------------|---------------------|-------|----|-----|--------|----|-----------------------|--------------------|
| C 100 | 16 | 2.7 | 302 | 7 | US-09-912-292-38493 | Sequence 38493, A | 172 | 16 | 2.7 | 1324 | 7 | US-09-850-118-1432 | Sequence 1432, Ap |
| C 101 | 16 | 2.7 | 306 | 7 | US-09-853-369-3588 | Sequence 3588, Ap | 173 | 16 | 2.7 | 1403 | 6 | US-09-760-466-560 | Sequence 560, App |
| C 102 | 16 | 2.7 | 311 | 7 | US-09-877-486-3861 | Sequence 3861, Ap | C 174 | 16 | 2.7 | 1692 | 1 | PCT-US01-08631-122519 | Sequence 22519, A |
| C 103 | 16 | 2.7 | 315 | 7 | US-09-867-716-6790 | Sequence 6790, Ap | C 175 | 16 | 2.7 | 1770 | 7 | US-09-902-540-4547 | Sequence 4547, App |
| C 104 | 16 | 2.7 | 328 | 8 | US-09-724-671-18201 | Sequence 18201, A | C 176 | 16 | 2.7 | 1975 | 6 | US-09-760-466-4547 | Sequence 472, App |
| C 105 | 16 | 2.7 | 331 | 7 | US-60-253-456-26975 | Sequence 26975, A | C 177 | 16 | 2.7 | 1992 | 1 | PCT-US01-08631-26331 | Sequence 26331, A |
| C 106 | 16 | 2.7 | 335 | 7 | US-09-724-671-4901 | Sequence 4901, Ap | C 178 | 16 | 2.7 | 2058 | 6 | US-09-758-469-1197 | Sequence 1197, App |
| C 107 | 16 | 2.7 | 337 | 8 | US-60-253-457-29711 | Sequence 29711, A | C 180 | 16 | 2.7 | 2140 | 8 | US-60-278-258-11997 | Sequence 11997, A |
| C 108 | 16 | 2.7 | 338 | 7 | US-09-724-671-18461 | Sequence 18461, A | C 181 | 16 | 2.7 | 2268 | 1 | PCT-US01-08631-119259 | Sequence 19259, A |
| C 109 | 16 | 2.7 | 346 | 7 | US-09-912-292-8604 | Sequence 8604, Ap | C 182 | 16 | 2.7 | 2327 | 8 | US-60-278-561-12029 | Sequence 12029, A |
| C 110 | 16 | 2.7 | 363 | 7 | US-09-724-671-18790 | Sequence 18790, A | C 183 | 16 | 2.7 | 2331 | 1 | PCT-US01-08631-122880 | Sequence 22880, A |
| C 111 | 16 | 2.7 | 371 | 6 | US-09-888-888-12081 | Sequence 12081, A | C 184 | 16 | 2.7 | 2345 | 8 | US-60-278-258-4998 | Sequence 4998, Ap |
| C 112 | 16 | 2.7 | 371 | 6 | US-09-866-555-19255 | Sequence 19255, A | C 185 | 16 | 2.7 | 2450 | 1 | PCT-US01-08631-120712 | Sequence 20712, A |
| C 113 | 16 | 2.7 | 372 | 8 | US-60-253-378-35457 | Sequence 35457, A | C 186 | 16 | 2.7 | 2550 | 7 | US-09-902-540-8526 | Sequence 8526, Ap |
| C 114 | 16 | 2.7 | 382 | 8 | US-60-252-833-13311 | Sequence 13311, A | C 187 | 16 | 2.7 | 2688 | 6 | US-09-598-075-154 | Sequence 154, App |
| C 115 | 16 | 2.7 | 387 | 7 | US-09-724-750-24153 | Sequence 24153, A | C 188 | 16 | 2.7 | 2688 | 6 | US-09-598-075-154 | Sequence 154, App |
| C 116 | 16 | 2.7 | 396 | 1 | PCT-US01-08631-116672 | Sequence 16672, A | C 189 | 16 | 2.7 | 3331 | 6 | US-09-235-810-31 | Sequence 31, App1 |
| C 117 | 16 | 2.7 | 397 | 6 | US-09-904-703-2778 | Sequence 2778, A | C 190 | 16 | 2.7 | 3331 | 6 | US-09-373-658-31 | Sequence 31, App1 |
| C 118 | 16 | 2.7 | 401 | 7 | US-09-850-147-3027 | Sequence 3027, Ap | C 191 | 16 | 2.7 | 3331 | 6 | US-60-171-503-31 | Sequence 31, App1 |
| C 119 | 16 | 2.7 | 409 | 8 | US-60-253-654-23916 | Sequence 23916, A | C 192 | 16 | 2.7 | 3342 | 8 | US-60-278-258-11917 | Sequence 11917, A |
| C 120 | 16 | 2.7 | 409 | 8 | US-60-255-592-23916 | Sequence 23916, A | C 193 | 16 | 2.7 | 3429 | 1 | PCT-US01-08631-129177 | Sequence 29177, A |
| C 121 | 16 | 2.7 | 410 | 8 | US-60-253-456-7084 | Sequence 7084, Ap | C 194 | 16 | 2.7 | 3684 | 8 | US-60-278-258-6146 | Sequence 6146, Ap |
| C 122 | 16 | 2.7 | 421 | 7 | US-09-692-257A-6018 | Sequence 6018, Ap | C 195 | 16 | 2.7 | 3718 | 6 | US-09-620-3128-74 | Sequence 74, App1 |
| C 123 | 16 | 2.7 | 432 | 7 | US-09-724-750-22736 | Sequence 22736, A | C 196 | 16 | 2.7 | 4369 | 6 | US-09-764-891-8870 | Sequence 8870, App |
| C 124 | 16 | 2.7 | 435 | 7 | US-09-724-750-21484 | Sequence 21484, A | C 197 | 16 | 2.7 | 4641 | 1 | PCT-US01-08631-125887 | Sequence 25887, A |
| C 125 | 16 | 2.7 | 427 | 6 | US-09-667-188A-11299 | Sequence 11299, Ap | C 198 | 16 | 2.7 | 7213 | 7 | US-60-278-258-39933 | Sequence 39933, A |
| C 126 | 16 | 2.7 | 431 | 6 | US-09-909-629-35857 | Sequence 35857, A | C 199 | 16 | 2.7 | 7671 | 8 | US-60-278-232-8827 | Sequence 8827, Ap |
| C 127 | 16 | 2.7 | 437 | 7 | US-09-773-476-1127 | Sequence 127, App | C 200 | 16 | 2.7 | 8046 | 1 | PCT-US01-08631-129175 | Sequence 29175, A |
| C 128 | 16 | 2.7 | 441 | 7 | US-09-834-366-36358 | Sequence 36358, A | C 201 | 16 | 2.7 | 9321 | 7 | US-09-902-540-898 | Sequence 898, App |
| C 129 | 16 | 2.7 | 443 | 6 | US-09-909-629-37049 | Sequence 37049, A | C 202 | 16 | 2.7 | 15172 | 7 | US-09-902-540-1086 | Sequence 1086, App |
| C 130 | 16 | 2.7 | 446 | 6 | US-09-909-629-32626 | Sequence 32626, A | C 203 | 16 | 2.7 | 18339 | 7 | US-09-881-797-808 | Sequence 808, App |
| C 131 | 16 | 2.7 | 451 | 8 | US-60-253-654-23636 | Sequence 23636, A | C 204 | 16 | 2.7 | 19222 | 7 | US-09-902-540-11192 | Sequence 1192, Ap |
| C 132 | 16 | 2.7 | 451 | 8 | US-60-255-592-23636 | Sequence 23636, A | C 205 | 16 | 2.7 | 21988 | 6 | US-09-803-736-347 | Sequence 347, App |
| C 133 | 16 | 2.7 | 461 | 7 | US-09-834-366-12022 | Sequence 12022, A | C 206 | 16 | 2.7 | 24757 | 6 | US-09-764-868-1347 | Sequence 1347, Ap |
| C 134 | 16 | 2.7 | 473 | 7 | US-09-864-761-2646 | Sequence 2646, Ap | C 207 | 16 | 2.7 | 25497 | 7 | US-09-902-540-1224 | Sequence 1224, Ap |
| C 135 | 16 | 2.7 | 475 | 8 | US-60-253-652-14488 | Sequence 14488, A | C 208 | 16 | 2.7 | 29899 | 7 | US-09-902-540-1265 | Sequence 1265, Ap |
| C 136 | 16 | 2.7 | 476 | 8 | US-60-253-653-3489 | Sequence 3489, A | C 209 | 16 | 2.7 | 39678 | 7 | US-09-764-905-39808 | Sequence 39808, A |
| C 137 | 16 | 2.7 | 482 | 7 | US-09-864-761-10572 | Sequence 10572, A | C 210 | 16 | 2.7 | 41927 | 7 | US-09-902-540-11268 | Sequence 11268, A |
| C 138 | 16 | 2.7 | 486 | 7 | US-09-824-518-5079 | Sequence 5079, Ap | C 211 | 16 | 2.7 | 76304 | 8 | US-09-289-426-1 | Sequence 1, App1 |
| C 139 | 16 | 2.7 | 487 | 7 | US-09-864-761-5298 | Sequence 5298, Ap | C 212 | 16 | 2.7 | 81662 | 6 | US-09-803-736-951 | Sequence 951, App1 |
| C 140 | 16 | 2.7 | 488 | 7 | US-09-912-292-43627 | Sequence 43627, A | C 213 | 16 | 2.7 | 84702 | 6 | US-09-803-736-474 | Sequence 474, App |
| C 141 | 16 | 2.7 | 491 | 8 | US-60-252-833-13471 | Sequence 13471, A | C 214 | 16 | 2.7 | 87210 | 6 | US-09-803-736-767 | Sequence 767, App |
| C 142 | 16 | 2.7 | 497 | 6 | US-09-866-555-6172 | Sequence 6172, Ap | C 215 | 16 | 2.7 | 99547 | 6 | US-09-803-736-1121 | Sequence 1121, App |
| C 143 | 16 | 2.7 | 501 | 6 | US-09-796-692-4449 | Sequence 4449, Ap | C 216 | 16 | 2.7 | 101647 | 6 | US-09-803-736-380 | Sequence 380, App |
| C 144 | 16 | 2.7 | 503 | 6 | US-09-796-692-5825 | Sequence 5825, Ap | C 217 | 16 | 2.7 | 114505 | 6 | US-09-803-736-129 | Sequence 129, App |
| C 145 | 16 | 2.7 | 507 | 7 | US-09-840-145-11795 | Sequence 11795, A | C 218 | 16 | 2.7 | 145744 | 6 | US-09-803-736-1465 | Sequence 1465, Ap |
| C 146 | 16 | 2.7 | 525 | 7 | US-09-902-540-9559 | Sequence 9559, Ap | C 219 | 16 | 2.7 | 165200 | 6 | US-09-803-110-205 | Sequence 205, App |
| C 147 | 16 | 2.7 | 553 | 6 | US-09-804-730-6011 | Sequence 6011, Ap | C 220 | 16 | 2.7 | 186449 | 7 | US-09-881-797-715 | Sequence 715, App |
| C 148 | 16 | 2.7 | 556 | 7 | US-09-861-478-273 | Sequence 273, App | C 221 | 16 | 2.7 | 199268 | 7 | US-09-881-797-697 | Sequence 697, App |
| C 149 | 16 | 2.7 | 558 | 7 | US-09-861-478-29 | Sequence 29, App1 | C 222 | 16 | 2.7 | 342340 | 6 | US-09-803-110-219 | Sequence 219, App |
| C 150 | 16 | 2.7 | 601 | 6 | US-09-758-450-458 | Sequence 458, App | C 223 | 16 | 2.7 | 51 | 8 | US-60-278-232-1843 | Sequence 1843, Ap |
| C 151 | 16 | 2.7 | 614 | 8 | US-60-253-378-36331 | Sequence 36331, A | C 224 | 15 | 2.5 | 57 | 5 | US-60-278-232-3958 | Sequence 3958, Ap |
| C 152 | 16 | 2.7 | 628 | 8 | US-60-253-652-11254 | Sequence 11254, A | C 225 | 15 | 2.5 | 51 | 8 | US-60-278-232-3958 | Sequence 3958, Ap |
| C 153 | 16 | 2.7 | 634 | 8 | US-09-834-366-10416 | Sequence 10416, A | C 226 | 15 | 2.5 | 65 | 7 | US-09-908-975-28318 | Sequence 28318, A |
| C 154 | 16 | 2.7 | 651 | 6 | US-09-738-626-982 | Sequence 982, App | C 227 | 15 | 2.5 | 65 | 8 | US-60-287-724-28318 | Sequence 28318, A |
| C 155 | 16 | 2.7 | 713 | 7 | US-09-861-478-316 | Sequence 316, App | C 228 | 15 | 2.5 | 106 | 7 | US-09-834-366-33225 | Sequence 33225, A |
| C 156 | 16 | 2.7 | 756 | 7 | US-09-902-540-5475 | Sequence 5475, App | C 229 | 15 | 2.5 | 155 | 18 | US-09-692-257A-2050 | Sequence 2050, App |
| C 157 | 16 | 2.7 | 876 | 6 | US-09-760-451-93 | Sequence 93, App1 | C 230 | 15 | 2.5 | 155 | 8 | US-60-253-457-40041 | Sequence 40041, A |
| C 158 | 16 | 2.7 | 909 | 6 | US-09-758-474-223 | Sequence 223, App1 | C 231 | 15 | 2.5 | 213 | 4 | US-08-804-730-21002 | Sequence 21002, A |
| C 159 | 16 | 2.7 | 919 | 6 | US-09-760-451-45 | Sequence 45, App1 | C 232 | 15 | 2.5 | 213 | 4 | US-08-798-0748-605 | Sequence 605, App |
| C 160 | 16 | 2.7 | 930 | 7 | US-09-902-540-5570 | Sequence 5570, App | C 233 | 15 | 2.5 | 213 | 4 | US-08-798-0748-605 | Sequence 605, App |
| C 161 | 16 | 2.7 | 976 | 6 | US-09-760-466-85 | Sequence 85, App1 | C 234 | 15 | 2.5 | 213 | 4 | US-08-798-0748-605 | Sequence 605, App |
| C 162 | 16 | 2.7 | 1030 | 7 | US-09-902-540-3902 | Sequence 3902, Ap | C 235 | 15 | 2.5 | 222 | 7 | US-09-840-213-38267 | Sequence 38267, A |
| C 163 | 16 | 2.7 | 1038 | 6 | US-09-803-110-6914 | Sequence 6914, Ap | C 236 | 15 | 2.5 | 222 | 7 | US-09-877-486-4991 | Sequence 4991, Ap |
| C 164 | 16 | 2.7 | 1070 | 7 | US-09-867-716-17316 | Sequence 17316, A | C 237 | 15 | 2.5 | 224 | 7 | US-09-877-486-2809 | Sequence 2809, Ap |
| C 165 | 16 | 2.7 | 1107 | 8 | US-60-309-870-59 | Sequence 870, App1 | C 238 | 15 | 2.5 | 224 | 7 | US-09-877-486-7079 | Sequence 7079, App |
| C 166 | 16 | 2.7 | 1146 | 6 | US-09-803-110-2857 | Sequence 2857, App1 | C 239 | 15 | 2.5 | 225 | 5 | US-09-824-559-3218 | Sequence 3218, Ap |
| C 167 | 16 | 2.7 | 1178 | 7 | US-09-881-797-4070 | Sequence 4070, App | C 240 | 15 | 2.5 | 226 | 8 | US-09-877-486-588 | Sequence 588, App |
| C 168 | 16 | 2.7 | 1236 | 1 | PCT-US01-16450-742 | Sequence 742, App | C 241 | 15 | 2.5 | 236 | 8 | US-60-253-378-36188 | Sequence 36188, A |
| C 169 | 16 | 2.7 | 1238 | 8 | US-60-309-870-58 | Sequence 58, App1 | C 242 | 15 | 2.5 | 244 | 7 | US-09-853-369-5124 | Sequence 5124, Ap |
| C 170 | 16 | 2.7 | 1266 | 8 | US-60-279-526-251 | Sequence 251, App | C 243 | 15 | 2.5 | 246 | 7 | US-09-724-671-20713 | Sequence 20713, A |
| C 171 | 16 | 2.7 | 1267 | 1 | PCT-US01-16450-1306 | Sequence 1306, App | C 244 | 15 | 2.5 | 249 | 6 | US-09-758-470-98 | Sequence 98, App1 |

| | | | | | | | |
|-------|-----|-----|---------|----|------------------|----------------------|-------------------|
| C 975 | 15 | 2.5 | 3309400 | 6 | US-09-7378-826-1 | Sequence 1, Appl | |
| 576 | -14 | 2.4 | 51 | 20 | 6 | US-09-1987-4524-6030 | Sequence 6030, Ap |
| 577 | -14 | 2.4 | 51 | 8 | 6 | US-60-2278-258-1803 | Sequence 1803, Ap |
| C 978 | -14 | 2.4 | 51 | 8 | 6 | US-60-278-258-1803 | Sequence 1803, Ap |
| C 979 | -14 | 2.4 | 51 | 8 | 6 | US-60-278-258-1803 | Sequence 3020, Ap |
| 980 | -14 | 2.4 | 51 | 8 | 6 | US-60-278-561-1706 | Sequence 8103, Ap |
| C 981 | -14 | 2.4 | 51 | 8 | 6 | US-60-278-561-12127 | Sequence 7706, Ap |
| C 982 | -14 | 2.4 | 51 | 8 | 6 | US-60-278-232-2159 | Sequence 12127, A |
| 983 | -14 | 2.4 | 51 | 8 | 6 | US-60-278-232-2159 | Sequence 2159, Ap |
| C 984 | -14 | 2.4 | 60 | 7 | 7 | US-09-908-975-14472 | Sequence 11815, A |
| C 985 | -14 | 2.4 | 60 | 7 | 7 | US-09-908-975-14472 | Sequence 14472, A |
| C 986 | -14 | 2.4 | 60 | 7 | 7 | US-60-287-724-14472 | Sequence 20597, A |
| C 987 | -14 | 2.4 | 60 | 8 | 8 | US-60-287-724-20597 | Sequence 14472, A |
| C 988 | -14 | 2.4 | 63 | 8 | 8 | US-60-253-378-12103 | Sequence 20597, A |
| 989 | -14 | 2.4 | 65 | 7 | 7 | US-09-908-975-26346 | Sequence 12103, A |
| 990 | -14 | 2.4 | 65 | 8 | 8 | US-60-287-724-26346 | Sequence 26346, A |
| 991 | -14 | 2.4 | 71 | 8 | 8 | US-60-253-654-18952 | Sequence 26346, A |
| 992 | -14 | 2.4 | 71 | 8 | 8 | US-60-253-592-18952 | Sequence 18952, A |
| 993 | -14 | 2.4 | 91 | 8 | 8 | US-60-253-654-29184 | Sequence 18952, A |
| 994 | -14 | 2.4 | 91 | 8 | 8 | US-60-253-592-29184 | Sequence 29184, A |
| 995 | -14 | 2.4 | 96 | 8 | 8 | US-60-253-654-7998 | Sequence 7998, A |
| 996 | -14 | 2.4 | 98 | 7 | 7 | US-60-253-592-7998 | Sequence 7998, Ap |
| 997 | -14 | 2.4 | 98 | 7 | 7 | US-09-834-366-33476 | Sequence 33476, A |
| 998 | -14 | 2.4 | 103 | 7 | 7 | US-09-864-761-26081 | Sequence 33476, A |
| C 999 | -14 | 2.4 | 108 | 7 | 7 | US-09-917-292-20683 | Sequence 26081, A |
| 1000 | -14 | 2.4 | 113 | 6 | 6 | US-09-688-848-3173 | Sequence 20683, A |
| | | | | | | Sequence 3173, Ap | |

ALIGNMENTS

```

RESULT 1
US-09-138-735-1
: Sequence 1, Application US/09138735
: GENERAL INFORMATION:
: APPLICANT: PARANHOS-BACCALA, Glaucia
: APPLICANT: LESENECHAL, Mylene
: APPLICANT: JOUIVER, Michel
: TITLE OF INVENTION: TRYPANOSOMA CRUZI ANTIGEN, GENE ENCODING THEREFOR AND METHODS OF
: TITLE OF INVENTION: DETECTING AND TREATING CHAGAS DISEASE
: FILE REFERENCE: WPI 36400B
: CURRENT APPLICATION NUMBER: US/09/138,735
: CURRENT FILING DATE: 1998-08-24
: PRIOR APPLICATION NUMBER: US 08/480,917
: PRIOR FILING DATE: 1995-06-07
: PRIOR APPLICATION NUMBER: FR 94/10132
: PRIOR FILING DATE: 1994-08-12
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 1
: LENGTH: 3402
: TYPE: DNA
: ORGANISM: Trypanosoma cruzi
: US-09-138-735-1

```

| Query Match | Best Local Similarity | Score 594: | DB 6: | Length 3402: |
|--|---|------------|-------|--------------|
| Matches 594; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | 100.0%; | 100.0%; | | |
| QY 1 | caggtacagcgttaacgcgctttgtcttcatacgttaacgcgcgaacgtatgctgtctctgtgc | 60 | | |
| Db 1232 | caggtacagcgttaacgcgctttgtcttcatacgttaacgcgcgaacgtatgctgtctctgtgc | 1291 | | |
| QY 61 | gacgtgtcgtatcgtattgacgacatcttcatacttcggtctctcccgcaaggagaagaacagcag | 120 | | |
| Db 1292 | gacgtgtcgtatcgtattgacgacatcttcatacttcggtctctcccgcaaggagaagaagaacagcag | 1351 | | |
| QY 121 | ccaaagccaaaataacatcgtatgtgcgcgcgcgcgaacccgggggtgtgtctccctcgggcact | 180 | | |
| Db 1352 | ccaaagccaaaataacatcgtatgtgcgcgcgcgcgaacccgggggtgtgtctccctcgggcact | 1411 | | |
| QY 181 | gaacgcgcgcagtagcagtcataccaatacgaattctgcgcgtctgtcgtacatccctcgtacata | 240 | | |

| | | | |
|----|------|--|------|
| Db | 1412 | gagcgcgcggtatgcagttcattaccataagactctcgcgcgtcgatccccctgaccca | 1471 |
| QY | 241 | ccccctgttccaagccagccaaagcagccgcctctctgcgcgcgcgacatcgctgag | 300 |
| Db | 1472 | ccccctttccaagccagccaaagcagccgcctctcgcgcgcgcgacatcgctgag | 1531 |
| QY | 301 | ccgcacgcctgggggagcaagatcatctgttaactctgataccaactgggataatgtaacc | 360 |
| Db | 1532 | ccgcacgcctgggggagcaagatcatctgttaactctgataccaactgggataatgtaacc | 1591 |
| QY | 361 | caaaagagcgtctcgtlcaagcacttgagcgccgcgcgcagacagagtlctaagcgagtgaactcc | 420 |
| Db | 1592 | caaaagagcgtctcgtlcaagcacttgagcgccgcgcgcagacagagtlctaagcgagtgaactcc | 1651 |
| QY | 421 | acgactaccgcgccccgcgcgcgcgaacaaatgcataagggacaacatggccgcgaactgtgaacgct | 480 |
| Db | 1652 | acgactaccgcgccccgcgcgcgcgaacaaatgcataagggacaacatggccgcgaactgtgaacgct | 1711 |
| QY | 481 | ggacttggctggcagcttaataatggctggccagcgccgactctctcccaacagccgcgcgaaga | 540 |
| Db | 1712 | ggacttggctggcagcttaataatggctggccagcgccgactctctcccaacagccgcgcgaaga | 1771 |
| QY | 541 | ccaacagagaagaagaagaagcctccgcgcgcatctggaacagactctcgctggcgata | 594 |
| Db | 1772 | ccaacagagaagaagaagaagcctccgcgcgcatctggaacagactctcgctggcgata | 1825 |

2

```

GENERAL INFORMATION:
APPLICANT: PARANHOS-BACCALA, Glaucia
APPLICANT: LESENECHAL, Mylene
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: TYPYANOSOMA CRUZI ANTIGEN, GENE ENCODING THEREFOR AND METHODS
TITLE OF INVENTION: DETECTING AND TREATING CHAGAS DISEASE
FILE REFERENCE: WPI 36400B
CURRENT APPLICATION NUMBER: US/09/138,735
CURRENT FILING DATE: 1998-08-24
PRIOR APPLICATION NUMBER: US 08/480,917
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: FR 94/10132
PRIOR FILING DATE: 1994-08-12
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn version 3.0
SEQ ID NO 9
LENGTH: 22
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: probe/primer
US-09-138-735-9

```

```

Query Match          3.7%; Score 22; DB 6;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      35  cagccgacgctagctgcgtcct 56
          |||||
Db       1  cagccgacgctagctgcgtcct 22

RESULT      3
US-09-803-736-19/c
: Sequence 19, Application US/09803736
: GENERAL INFORMATION:
: APPLICANT: Bush, David F.
: APPLICANT: Levin, Irena M.
: APPLICANT: Norris, Susan R.
: APPLICANT: Rounsley, Steven D.
: APPLICANT: Wiegand, Roger C.

```

TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
FILE REFERENCE: 38-10(15493)D
CURRENT APPLICATION NUMBER: US/09/803,736
CURRENT FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: US 09/534,859
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: Identified by Attorney Docket number 04983.0206CPUS01 38-10
PRIOR FILING DATE: 2000-10-20
NUMBER OF SEQ ID NOS: 1582
SEQ ID NO 19
LENGTH: 91470
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-803-736-19

Query Match 3.4%; Score 20; DB 6; Length 91470;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 212 cttctgcgcgtcgtcgtcc 231
|||||
Db 42747 CTTCTGCCGCTGCTGCATCC 42728

RESULT 4
US-09-138-735-5
Sequence 5, Application US/09138735
GENERAL INFORMATION:
APPLICANT: PARANHOS-BACCALA, Glaucia
APPLICANT: LESENECHAL, Mylene
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: TRYPANOSOMA CRUZI ANTIGEN, GENE ENCODING THEREFOR AND METHODS OF
FILE REFERENCE: WPB 36400B
CURRENT APPLICATION NUMBER: US/09/138,735
CURRENT FILING DATE: 1998-08-24
PRIOR APPLICATION NUMBER: US 08/480,917
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: FR 94/10132
PRIOR FILING DATE: 1994-08-12
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 18
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: primer
US-09-138-735-5

Query Match 3.0%; Score 18; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 172 tcgggcactgacgcggcg 189
|||||
Db 1 tcgggcactgacgcggcg 18

RESULT 5
US-09-138-735-8/c
Sequence 8, Application US/09138735
GENERAL INFORMATION:
APPLICANT: PARANHOS-BACCALA, Glaucia
APPLICANT: LESENECHAL, Mylene
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: TRYPANOSOMA CRUZI ANTIGEN, GENE ENCODING THEREFOR AND METHODS OF
FILE REFERENCE: WPB 36400B
CURRENT APPLICATION NUMBER: US/09/138,735
CURRENT FILING DATE: 1998-08-24

PRIOR APPLICATION NUMBER: US 08/480,917
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: FR 94/10132
PRIOR FILING DATE: 1994-08-12
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 18
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: probe/primer
US-09-138-735-8

Query Match 3.0%; Score 18; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 actctgcgcgtcgtcga 228
|||||
Db 18 ACTTCTGCCGCTGCTGCA 1

RESULT 6
US-09-866-555-4980/C
Sequence 4980, Application US/09866555
GENERAL INFORMATION:
APPLICANT: Harlocker, Susan L.
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Algate, Paul A.
APPLICANT: Jones, Robert
APPLICANT: Klee, Jennifer
TITLE OF INVENTION: LUNG TUMOR-ASSOCIATED SEQUENCES AND METHODS OF USE
FILE REFERENCE: 210121.580
CURRENT APPLICATION NUMBER: US/09/866,555
CURRENT FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 20487
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4980
LENGTH: 261
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)..(261)
OTHER INFORMATION: n = A,T,C or G
US-09-866-555-4980

Query Match 3.0%; Score 18; DB 6; Length 261;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 gtagctgcgtcgtcgtg 61
|||||
Db 18 GTAGCTGCCTCTGCTG 1

RESULT 7
US-09-863-524-3874
Sequence 3874, Application US/09863524
GENERAL INFORMATION:
APPLICANT: Ito, Laura Y.
APPLICANT: Gooding, Douglas H.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN ROOTS
FILE REFERENCE: PI-0003 US
CURRENT APPLICATION NUMBER: US/09/863,524
CURRENT FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: 60/077,831
PRIOR FILING DATE: 1998-03-13

```

; NUMBER OF SEQ ID NOS: 6814
; SOFTWARE: PERL Program
; SEQ ID NO: 3874
; LENGTH: 297
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 700047828h1
US-09-863-524-3874

Query Match          3.0%; Score 18; DB 7; Length 297;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 386 cgcgcgcgcacgacgagt 403
    |||
Db 247 cgcgcgcgcacgacgagt 264

RESULT 8
US-09-764-891-1027/C
; Sequence 1027, Application US/09764891
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO: 1027
; LENGTH: 328
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-1027

Query Match          3.0%; Score 18; DB 6; Length 328;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 actctcgcgcgtcgtcga 228
    |||
Db 251 ACTTCGTCCGCGCTGTGCA 234

RESULT 9
US-09-540-213-5492
; Sequence 5492, Application US/09540213
; GENERAL INFORMATION:
; APPLICANT: Sellhammer, Jeffrey J.
; APPLICANT: Deleagane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF LIVER AND PANCREAS
; FILE REFERENCE: PD-1031 CIP
; CURRENT APPLICATION NUMBER: US/09/540,213
; CURRENT FILING DATE: 2000-03-31
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 53844
; SEQ ID NO: 5492
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00532370
; NAME/KEY: unsure
```

```

; LOCATION: 4, 8, 304, 307
; OTHER INFORMATION: a, t, c, g, or other
US-09-540-213-5492

Query Match          3.0%; Score 18; DB 7; Length 348;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 cccctgcatacccccctg 247
    |||
Db 29 cccctgcatacccccctg 46

RESULT 10
US-09-667-188A-4578/C
; Sequence 4578, Application US/09667188A
; GENERAL INFORMATION:
; APPLICANT: Andersen, Scott E.
; APPLICANT: Masucci, James D.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51464)B
; CURRENT APPLICATION NUMBER: US/09/667,188A
; CURRENT FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/155,006
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 8135
; SEQ ID NO: 4578
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Trillium aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3281-014-pl-K1-E7
US-09-667-188A-4578

Query Match          3.0%; Score 18; DB 6; Length 371;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 274 cctctcgcgcgcgcgcga 291
    |||
Db 153 CCTCTGTCGCGCGCGCGCA 136

RESULT 11
US-09-864-761-52/C
; Sequence 52, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
```



```
; OTHER INFORMATION: n = A,T,C or G
US-09-904-809-19378

Query Match          3.0%; Score 18; DB 7; Length 468;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 cccctgcataccccctg 247
      |||||
Db 393 CCCCTGCATCACCCCTG 376

RESULT 12
US-09-904-809-19378/c
; Sequence 19378, Application US/09904809
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-757
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 09/234,611
; PRIOR FILING DATE: 1999-01-22
; NUMBER OF SEQ ID NOS: 21025
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19378
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(476)
```

```
; OTHER INFORMATION: n = A,T,C or G
US-09-904-809-19378

Query Match          3.0%; Score 18; DB 6; Length 476;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 gtagctgcctctgctg 61
      |||||
Db 384 GTAGCTGCCTCTGCTG 367

RESULT 13
US-09-909-627-11623/c
; Sequence 11623, Application US/09909627
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-766
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 09/277,227
; PRIOR FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 23680
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11623
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(476)
; OTHER INFORMATION: n = A,T,C or G
US-09-909-627-11623

Query Match          3.0%; Score 18; DB 6; Length 476;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 gtagctgcctctgctg 61
      |||||
Db 384 GTAGCTGCCTCTGCTG 367

RESULT 14
US-09-864-761-16893/c
; Sequence 16893, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
```

```
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 16893
;; LENGTH: 478
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC005921.3
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.9
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6.9
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
;; OTHER INFORMATION: SWISSPROT HIT: P50555, EVALUATE 5.60e-01
;; OTHER INFORMATION: EST_HUMAN HIT: BE74137.1, EVALUATE 0.00e+00
;; OTHER INFORMATION: NT HIT: X99000.1, EVALUATE 3.00e-01
;; US-09-864-761-16893

Query Match          3.0%; Score 18; DB 7; Length 478;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 230 cccctgcatcacccctg 247
    |||||||
Db 298 CCCCTGCATCACCCCTG 281

RESULT 15
US-09-840-145-13709/c
; Sequence 13709, Application US/09840145
; GENERAL INFORMATION:
; APPLICANT: Adams, et. al.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products
; FILE REFERENCE: POLAC1
; CURRENT APPLICATION NUMBER: US/09/840,145
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 08/276,163
; PRIOR FILING DATE: 1994-07-15
; NUMBER OF SEQ ID NOS: 15314
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13709
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (11)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (30)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (44)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (46)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (51)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (60)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (66)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (148)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (185)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (306)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (321)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (345)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (359)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (364)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (375)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (425)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (427)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (440)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (443)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (470)
;; OTHER INFORMATION: n equals a,t,g, or c
```


TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC004
CURRENT APPLICATION NUMBER: US/09/764,905
CURRENT FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/225,757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997

PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241,787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,245
PRIOR FILING DATE: 2000-11-17

;; PRIOR APPLICATION NUMBER: 60/249,244
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,217
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,211
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,215
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,264
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,214
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,297
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/232,400
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/231,242
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,081
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,080
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,414
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,244
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/233,064
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/233,063
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,397
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,399
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,401
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/241,808
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,826
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,786
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,221
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/246,475
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/231,243
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/233,065
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,398
;;
Query Match 3.0%; Score 18; DB 7; Length 15270;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 227 catccctgcataccccc 244
Db 11976 CATCCCTGCATCACCCC 11959
;;
RESULT 20
US-09-764-905-40566/c
;; Sequence 40566, Application US/09764905
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PC004
;; CURRENT APPLICATION NUMBER: US/09/764,905
;; PRIOR FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: 60/1179,065
;; PRIOR FILING DATE: 2000-01-31
;; PRIOR APPLICATION NUMBER: 60/180,628
;;

;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: 60/214,886
;; PRIOR FILING DATE: 2000-06-28
;; PRIOR APPLICATION NUMBER: 60/217,487
;; PRIOR FILING DATE: 2000-07-11
;; PRIOR APPLICATION NUMBER: 60/225,758
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/220,963
;; PRIOR FILING DATE: 2000-07-26
;; PRIOR APPLICATION NUMBER: 60/217,496
;; PRIOR FILING DATE: 2000-07-11
;; PRIOR APPLICATION NUMBER: 60/225,447
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/218,290
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/225,757
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/226,868
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/216,647
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/225,267
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/216,880
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/225,270
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/251,869
;; PRIOR FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/235,834
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/234,274
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: 60/234,223
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: 60/228,924
;; PRIOR FILING DATE: 2000-08-30
;; PRIOR APPLICATION NUMBER: 60/224,518
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/236,369
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/224,519
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/220,964
;; PRIOR FILING DATE: 2000-07-26
;; PRIOR APPLICATION NUMBER: 60/241,809
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/249,299
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/236,327
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/241,785
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/244,617
;; PRIOR FILING DATE: 2000-11-01
;; PRIOR APPLICATION NUMBER: 60/225,268
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/236,368
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/251,856
;; PRIOR FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/251,868
;; PRIOR FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/229,344
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/234,997
;; PRIOR FILING DATE: 2000-09-25
;; PRIOR APPLICATION NUMBER: 60/229,343
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,345
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,287
;; PRIOR FILING DATE: 2000-09-01
;;

PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241,787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,245
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,244
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,217
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,211
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,215

PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,264
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,214
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,297
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232,400
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231,242
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,081
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,414
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,244
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,064
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/233,063
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,397
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,399
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,401
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/241,808
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,826
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,786
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,221
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,475
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/231,243
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,065
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,398

Query Match 3.0%; Score 18; DB 7; Length 15270;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 227 catccctgcatcacc 244
DB 11976 CATCCCTGCATCACC 11959

RESULT 21
US-09-803-736-1397/C
Sequence 1397, Application US/09803736
GENERAL INFORMATION:
APPLICANT: Bush, David F.
APPLICANT: Levin, Irena M.
APPLICANT: Morris, Susan R.
APPLICANT: Rounsley, Steven D.
TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof.
FILE REFERENCE: 38-10(15493)D
CURRENT APPLICATION NUMBER: US/09/803,736
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: US 09/534,859
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: identified by Attorney Docket number 04983.0206CPU501 38
PRIOR FILING DATE: 2000-10-20
NUMBER OF SEQ ID NOS: 1582
SEQ ID NO 1397

LENGTH: 96424
TYPE: DNA
ORGANISM: Arabidopsis thaliana.
US-09-803-736-1397

Query Match 3.0%; Score 18; DB 6; Length 96424;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 202 accaatacgaactctgc 219
|||||
Db 81927 ACCAATACGACTCTGCC 81910

RESULT 22
US-09-724-671-17298
Sequence 17298, Application US/09724671
GENERAL INFORMATION:
APPLICANT: Watson, James D
APPLICANT: Murison, James G
TITLE OF INVENTION: polynucleotides, polypeptides expressed
by the polynucleotides and methods for their use.
FILE REFERENCE: 105002
CURRENT APPLICATION NUMBER: US/09/724, 671
CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: U.S. No. 60/171,678
PRIOR FILING DATE: 1999-12-23
NUMBER OF SEQ ID NOS: 21907
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17298
LENGTH: 241
TYPE: DNA
ORGANISM: Mouse
US-09-724-671-17298

Query Match 2.9%; Score 17; DB 7; Length 241;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 113 aacagcagccagccaa 129
|||||
Db 98 aacagcagccagccaa 114

RESULT 23
US-09-540-213-52685
Sequence 52685, Application US/09540213
GENERAL INFORMATION:
APPLICANT: Seilhamer, Jeffrey J.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Mullahy, Sara L.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES OF LIVER AND PANCREAS
FILE REFERENCE: PD-1031 CIP
CURRENT APPLICATION NUMBER: US/09/540,213
CURRENT FILING DATE: 2000-03-31
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 53844
SEQ ID NO 52685
LENGTH: 241
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID NO: rat00171041
US-09-540-213-52685

Query Match 2.9%; Score 17; DB 7; Length 241;

Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 113 aacagcagccagccaa 129
|||||
Db 71 aacagcagccagccaa 87

RESULT 24
US-09-540-213-44587
Sequence 44587, Application US/09540213
GENERAL INFORMATION:
APPLICANT: Seilhamer, Jeffrey J.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullahy, Sara J.
TITLE OF INVENTION: POLYNUCLEOTIDES OF LIVER AND PANCREAS
FILE REFERENCE: PD-1031 CIP
CURRENT APPLICATION NUMBER: US/09/540,213
CURRENT FILING DATE: 2000-03-31
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 53844
SEQ ID NO 44587
LENGTH: 264
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID NO: rat00163047
US-09-540-213-44587

Query Match 2.9%; Score 17; DB 7; Length 264;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 113 aacagcagccagccaa 129
|||||
Db 71 aacagcagccagccaa 87

RESULT 25
US-09-540-213-33702
Sequence 33702, Application US/09540213
GENERAL INFORMATION:
APPLICANT: Seilhamer, Jeffrey J.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullahy, Sara J.
TITLE OF INVENTION: POLYNUCLEOTIDES OF LIVER AND PANCREAS
FILE REFERENCE: PD-1031 CIP
CURRENT APPLICATION NUMBER: US/09/540,213
CURRENT FILING DATE: 2000-03-31
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 53844
SEQ ID NO 33702
LENGTH: 300
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID NO: rat00076798
US-09-540-213-33702

Query Match 2.9%; Score 17; DB 7; Length 300;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 aacagcagcagcgccaa 129
Db 76 aacagcagcagcgccaa 92

RESULT 26

US-09-724-671-8427
; Sequence 8427, Application US/09724671
; GENERAL INFORMATION:
; APPLICANT: Watson, James D
; APPLICANT: Murison, James G
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed
; TITLE OF INVENTION: by the polynucleotides and methods for their use.
; FILE REFERENCE: 105002
; CURRENT APPLICATION NUMBER: US/09/724,671
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: U.S. No. 60/171,678
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 21907
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8427
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Mouse
US-09-724-671-8427

Query Match 2.9%; Score 17; DB 7; Length 301;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 aacagcagcagcgccaa 129
Db 98 aacagcagcagcgccaa 114

RESULT 27

US-09-724-671-8448
; Sequence 8448, Application US/09724671
; GENERAL INFORMATION:
; APPLICANT: Watson, James D
; APPLICANT: Murison, James G
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed
; TITLE OF INVENTION: by the polynucleotides and methods for their use.
; FILE REFERENCE: 105002
; CURRENT APPLICATION NUMBER: US/09/724,671
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: U.S. No. 60/171,678
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 21907
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8448
; LENGTH: 303
; TYPE: DNA
; ORGANISM: Mouse
US-09-724-671-8448

Query Match 2.9%; Score 17; DB 7; Length 303;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 aacagcagcagcgccaa 129
Db 98 aacagcagcagcgccaa 114

RESULT 28

US-09-863-524-3604/C
; Sequence 3604, Application US/09863524
; GENERAL INFORMATION:
; APPLICANT: Ito, Laura Y.
; APPLICANT: Gooding, Douglas H.

; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN ROOTS
; FILE REFERENCE: PL-0003 US
; CURRENT APPLICATION NUMBER: US/09/863,524
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: 60/077,831
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 6814
; SOFTWARE: PERL Program
; SEQ ID NO 3604
; LENGTH: 322
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 700047418H1
; NAME/KEY: unsure
; LOCATION: 203, 238, 261, 276
; OTHER INFORMATION: a, t, c, g, or other
US-09-863-524-3604

Query Match 2.9%; Score 17; DB 7; Length 322;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 gtactgagcagcgccaa 155
Db 231 gtactgagcagcgccaa 215

RESULT 29

US-09-724-671-5748
; Sequence 5748, Application US/09724671
; GENERAL INFORMATION:
; APPLICANT: Watson, James D
; APPLICANT: Murison, James G
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed
; TITLE OF INVENTION: by the polynucleotides and methods for their use.
; FILE REFERENCE: 105002
; CURRENT APPLICATION NUMBER: US/09/724,671
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: U.S. No. 60/171,678
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 21907
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5748
; LENGTH: 353
; TYPE: DNA
; ORGANISM: Mouse
US-09-724-671-5748

Query Match 2.9%; Score 17; DB 7; Length 353;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 aacagcagcagcgccaa 129
Db 98 aacagcagcagcgccaa 114

RESULT 30

US-09-724-671-4945
; Sequence 4945, Application US/09724671
; GENERAL INFORMATION:
; APPLICANT: Watson, James D
; APPLICANT: Murison, James G
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed
; TITLE OF INVENTION: by the polynucleotides and methods for their use.
; FILE REFERENCE: 105002
; CURRENT APPLICATION NUMBER: US/09/724,671
; CURRENT FILING DATE: 2000-11-28

PRIOR APPLICATION NUMBER: U.S. No. 60/171,678
PRIOR FILING DATE: 1999-12-23
NUMBER OF SEQ ID NOS: 21907
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4945
LENGTH: 396
TYPE: DNA
ORGANISM: Mouse
US-09-724-671-4945

Query Match 2.9%; Score 17; DB 7; Length 396;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 113 aacagcagccagcgccaa 129
|||||
DB 98 aacagcagccagcgccaa 114

RESULT 31

US-09-804-730-15348
Sequence 15348, Application US/09804730
GENERAL INFORMATION:
APPLICANT: Andersen, Scott E.
APPLICANT: Hammond-Kosack, Kim
APPLICANT: Masucci, James D.
APPLICANT: Urban, Martin
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
FILE REFERENCE: 38-21(51838)B
CURRENT APPLICATION NUMBER: US/09/804,730
CURRENT FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: US 60/189,657
PRIOR FILING DATE: 2000-03-15
NUMBER OF SEQ ID NOS: 22828
SEQ ID NO 15348
LENGTH: 398
TYPE: DNA
ORGANISM: Trifolium aestivum
OTHER INFORMATION: Clone ID: LIB3399-028-P1-K1-B7
US-09-804-730-15348

Query Match 2.9%; Score 17; DB 6; Length 398;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 386 cgcgcgcagcagcagcagc 402
|||||
DB 270 cgcgcgcagcagcagcagc 286

RESULT 32

US-09-904-703-4316
Sequence 4316, Application US/09904703
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-758CON1
CURRENT APPLICATION NUMBER: US/09/904,703
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: 09/210,298
PRIOR FILING DATE: 1998-12-09
NUMBER OF SEQ ID NOS: 17812
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4316
LENGTH: 402
TYPE: DNA
ORGANISM: Homo sapiens
US-09-904-703-4316

Query Match 2.9%; Score 17; DB 6; Length 402;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 113 aacagcagccagcgccaa 129
|||||
DB 73 aacagcagccagcgccaa 89

RESULT 33

US-09-866-555-6918/C
Sequence 6918, Application US/09866555
GENERAL INFORMATION:
APPLICANT: Harlocker, Susan L.
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Algate, Paul A.
APPLICANT: Jones, Robert
APPLICANT: Klee, Jennifer
TITLE OF INVENTION: LUNG TUMOR-ASSOCIATED SEQUENCES AND METHODS OF USE
FILE REFERENCE: 210121.580
CURRENT APPLICATION NUMBER: US/09/866,555
CURRENT FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 20487
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6918
LENGTH: 404
TYPE: DNA
ORGANISM: Homo sapien
US-09-866-555-6918

Query Match 2.9%; Score 17; DB 6; Length 404;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 352 aatgtaccccaaggag 368
|||||
DB 326 AATGTACCCCAAGGAG 310

RESULT 34

US-09-692-257A-6664/C
Sequence 6664, Application US/09692257A
GENERAL INFORMATION:
APPLICANT: Miller, Phillip W.
APPLICANT: Peng, Ming
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
FILE REFERENCE: 38-21(15771)B
CURRENT APPLICATION NUMBER: US/09/692,257A
CURRENT FILING DATE: 2000-10-19
PRIOR APPLICATION NUMBER: US 60/162,747
PRIOR FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 14882
SEQ ID NO 6664
LENGTH: 419
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3137-020-Q1-K1-B5
US-09-692-257A-6664

Query Match 2.9%; Score 17; DB 7; Length 419;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 274 cctctgcgcgcgcgcgc 290
|||||
DB 284 CCTCTGCCGCGCGCGC 268

```
RESULT 35
US-09-804-730-8033
; Sequence 8033, Application US/09804730
; GENERAL INFORMATION:
; APPLICANT: Andersen, Scott E.
; APPLICANT: Hammond, Kosack, Kim
; APPLICANT: Masucci, James D.
; APPLICANT: Urban, Martin
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(51838)B
; CURRENT APPLICATION NUMBER: US/09/804,730
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: US 60/189,657
; PRIOR FILING DATE: 2000-03-15
; NUMBER OF SEQ ID NOS: 22828
; SEQ ID NO 8033
; LENGTH: 426
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: (1)..(426)
; LOCATION: (1)..(426)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB351-013-Pl-K1-A4
US-09-804-730-8033
```

```
Query Match          2.9%; Score 17; DB 6; Length 426;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 386 cgcgcgcacacgagcagg 402
      |||||
Db 117 cgcgcgcacacgagcagg 133
```

```
RESULT 36
US-09-904-809-4385/C
; Sequence 4385, Application US/09904809
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-757
; CURRENT APPLICATION NUMBER: US/09/904,809
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 09/234,611
; PRIOR FILING DATE: 1999-01-22
; NUMBER OF SEQ ID NOS: 21025
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4385
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)..(430)
; OTHER INFORMATION: n = A,T,C or G
US-09-904-809-4385
```

```
Query Match          2.9%; Score 17; DB 6; Length 430;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 515 cctcgtctccacagcc 531
      |||||
Db 173 cctcgtctccacagcc 157
```

```
RESULT 37
```

```
US-09-881-797-3357/C
; Sequence 3357, Application US/09881797
; GENERAL INFORMATION:
; APPLICANT: Tim Keith
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO RESPIRATORY DISEASES AND OBESITY
; FILE REFERENCE: H001-04
; CURRENT APPLICATION NUMBER: US/09/881,797
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/211,749
; PRIOR FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 4687
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3357
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Human
US-09-881-797-3357
```

```
Query Match          2.9%; Score 17; DB 7; Length 439;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 546 agcagaagaagcct 562
      |||||
Db 157 AGCAGAAGAAGCCT 141
```

```
RESULT 38
US-09-724-671-4408
; Sequence 4408, Application US/09724671
; GENERAL INFORMATION:
; APPLICANT: Watson, James D
; APPLICANT: Murison, James G
; TITLE OF INVENTION: polynucleotides, polypeptides expressed
; TITLE OF INVENTION: by the polynucleotides and methods for their use.
; FILE REFERENCE: 105002
; CURRENT APPLICATION NUMBER: US/09/724,671
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: U.S. No. 60/171,678
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 21907
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4408
; LENGTH: 444
; TYPE: DNA
; ORGANISM: Mouse
US-09-724-671-4408
```

```
Query Match          2.9%; Score 17; DB 7; Length 444;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 113 aacagcagcagccaa 129
      |||||
Db 98 aacagcagcagccaa 114
```

```
RESULT 39
US-09-758-463-219
; Sequence 219, Application US/09758463
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM011
; CURRENT APPLICATION NUMBER: US/09/758,463
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
```

NUMBER OF SEQ ID NOS: 1304
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 219
LENGTH: 468
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (390)
OTHER INFORMATION: n equals a,t,g, or c
US-09-758-463-219

Query Match 2.9%; Score 17; DB 6; Length 468;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 113 aacagcagcagcgccaa 129
|||||
Db 242 aacagcagcagcgccaa 258

RESULT 40
US-09-881-797-3356/c
Sequence 3356, Application US/09881797
GENERAL INFORMATION:
APPLICANT: Tim Keith
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES
FILE REFERENCE: HOM01-04
CURRENT APPLICATION NUMBER: US/09/881,797
CURRENT FILING DATE: 2001-06-14
PRIORITY APPLICATION NUMBER: 60/211,749
PRIOR FILING DATE: 2000-06-14
NUMBER OF SEQ ID NOS: 4687
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3356
LENGTH: 474
TYPE: DNA
ORGANISM: Human
US-09-881-797-3356

Query Match 2.9%; Score 17; DB 7; Length 474;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 546 aggaagaagaagcct 562
|||||
Db 190 AGGAGAGAAAGGCT 174

RESULT 41
US-60-253-652-24408/c
Sequence 24408, Application US/60253652
GENERAL INFORMATION:
APPLICANT: Glenn, Matthew R
TITLE OF INVENTION: Compositions Isolated from bovine
FILE REFERENCE: 1055p2
CURRENT APPLICATION NUMBER: US/60/253,652
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 29954
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 24408
LENGTH: 481
TYPE: DNA
ORGANISM: Bovine
US-60-253-652-24408

Query Match 2.9%; Score 17; DB 8; Length 481;

Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 483 atgtgtgcagctaata 499
|||||
Db 447 ATTGTTGCAGCTAATA 431

RESULT 42
US-09-823-241-6683
Sequence 6683, Application US/09823241
GENERAL INFORMATION:
APPLICANT: Gearling, David P.
APPLICANT: Holtzman, Douglas A.
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: 1600.2075-001
CURRENT APPLICATION NUMBER: US/09/823,241
CURRENT FILING DATE: 2001-03-30
PRIORITY APPLICATION NUMBER: US 60/193,434
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 10028
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6683
LENGTH: 492
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(492)
OTHER INFORMATION: n = A,T,C or G
US-09-823-241-6683

Query Match 2.9%; Score 17; DB 6; Length 492;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 113 aacagcagcagcgccaa 129
|||||
Db 78 aacagcagcagcgccaa 94

RESULT 43
US-09-850-147-15830
Sequence 15830, Application US/09850147
GENERAL INFORMATION:
APPLICANT: Andersen, Scott E.
APPLICANT: Edgerton, Michael D
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
FILE REFERENCE: 38-21(51914)B
CURRENT APPLICATION NUMBER: US/09/850,147
CURRENT FILING DATE: 2001-05-08
PRIORITY APPLICATION NUMBER: US 60/202,213
PRIOR FILING DATE: 2000-05-08
PRIORITY APPLICATION NUMBER: US 09/654,617
PRIOR FILING DATE: 2000-09-05
PRIORITY APPLICATION NUMBER: US 09/684,016
PRIOR FILING DATE: 2000-10-10
NUMBER OF SEQ ID NOS: 18014
SEQ ID NO 15830
LENGTH: 500
TYPE: DNA
ORGANISM: Sorghum bicolor
OTHER INFORMATION: Clone ID: LIB3481-018-P1-K1-A10
US-09-850-147-15830

Query Match 2.9%; Score 17; DB 7; Length 500;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 tttaagcgcagccaaag 264
|||||
Db 82 tttaagcgcagccaaag 98

RESULT 44
US-09-834-366-4311

; Sequence 4311, Application US/09834366
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Maline Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81 US2 REG
; CURRENT APPLICATION NUMBER: US/09/834,366
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 4311
; LENGTH: 501
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 121..501
US-09-834-366-4311

Query Match 2.9%; Score 17; DB 7; Length 501;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 aacagcagccagccaa 129
|||||
Db 126 aacagcagccagccaa 142

RESULT 45

US-09-758-463-239
; Sequence 239, Application US/09758463
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antipodides
; FILE REFERENCE: PM011
; CURRENT APPLICATION NUMBER: US/09/758,463
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 239
; LENGTH: 514
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-758-463-239

Query Match 2.9%; Score 17; DB 6; Length 514;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 aacagcagccagccaa 129
|||||
Db 137 aacagcagccagccaa 153

RESULT 46

US-09-821-837-4107
; Sequence 4107, Application US/09821837
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.2076-001
; CURRENT APPLICATION NUMBER: US/09/821,837
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 60/193,481
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 9928
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4107
; LENGTH: 562
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(562)
; OTHER INFORMATION: n = A,T,C or G
US-09-821-837-4107

Query Match 2.9%; Score 17; DB 7; Length 562;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 gcaagcgcgcctcctgc 281
|||||
Db 34 gcaagcgcgcctcctgc 50

RESULT 47

US-09-823-327-3606
; Sequence 3606, Application US/09823327
; GENERAL INFORMATION:
; APPLICANT: Richardson, Jennifer
; APPLICANT: Macbeth, Kyle J.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.2083-001
; CURRENT APPLICATION NUMBER: US/09/823,327
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,433
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 9746
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3606
; LENGTH: 643
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(643)
; OTHER INFORMATION: n = A,T,C or G
US-09-823-327-3606

Query Match 2.9%; Score 17; DB 7; Length 643;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 aacagcagccagccaa 129
|||||
Db 84 aacagcagccagccaa 100

RESULT 48

US-09-623-746A-60
; Sequence 60, Application US/09623746A
; GENERAL INFORMATION:

```
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM PROSTATE TISSUE
; FILE REFERENCE: ALBRE 10
; CURRENT APPLICATION NUMBER: US/09/623,746A
; PRIOR FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: PCT/DE99/00722
; PRIOR FILING DATE: 1999-03-09
; NUMBER OF SEQ ID NOS: 297
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
; LENGTH: 697
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-623-746A-60
```

```
Query Match          2.9%; Score 17; DB 6; Length 697;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 113 aacagcagccagcgccaa 129
    |||||
Db 110 aacagcagccagcgccaa 126
```

```
RESULT 49
PCT-US01-08631-9801/c
; Sequence 9801, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 9801
; LENGTH: 771
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (129)..(689)
; OTHER INFORMATION: 29% homologous to Santalum album proline rich
; OTHER INFORMATION: protein, accession number AF020261, Smith-Waterman Score=112.
PCT-US01-08631-9801
```

```
Query Match          2.9%; Score 17; DB 1; Length 771;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 257 cagccaagcagcgccg 273
    |||||
Db 602 CAGCCAGGACGCCGCG 586
```

```
RESULT 50
PCT-US01-08631-9802/c
; Sequence 9802, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
```

```
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 9802
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (8)..(493)
; OTHER INFORMATION: 30% homologous to Homo sapiens Human PRO1305 (UN0671) amino
; OTHER INFORMATION: acid sequence SEQ ID NO:153, accession number Y93386, Smith-Wate
PCT-US01-08631-9802
```

```
Query Match          2.9%; Score 17; DB 1; Length 781;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 257 cagccaagcagcgccg 273
    |||||
Db 557 CAGCCAGGACGCCGCG 541
```

```
Search completed: September 21, 2001, 13:33:10
Job time: 551 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 13:59:40 : Search time 15806.4 Seconds
(without alignments)
581.273 Million cell updates/sec

Title: US-09-138-735-1_COPY_1232_1825
Perfect score: 594
Sequence: 1 caggtacagcgtacagcgtc.....aaacagcctccgtgcgata 594

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1344157 seqs, 7733874588 residues

Word size: 0

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database:

GenEmbl:
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: gb_pl5:*
17: gb_ba2:*
18: gb_ba1:*
19: gb_ba2:*
20: gb_ba1:*
21: gb_ba2:*
22: gb_ba3:*
23: gb_in1:*
24: gb_in2:*
25: gb_in3:*
26: gb_om:*
27: gb_ov:*
28: gb_pat1:*
29: gb_pat2:*
30: gb_ph:*
31: gb_pl1:*
32: gb_pl2:*
33: gb_pl3:*
34: gb_pl4:*
35: gb_pl5:*
36: gb_ba2:*
37: gb_ba1:*
38: gb_ba2:*
39: gb_in1:*
40: gb_in2:*
41: gb_in3:*
42: gb_om:*
43: gb_ov:*

Oligo Search

5+

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 594 | 100.0 | 3402 | 9 | AR047920 |
| 2 | 492 | 82.8 | 3402 | 9 | A48910 |
| 3 | 492 | 82.8 | 3402 | 96 | TCU24190 |
| 4 | 22 | 3.7 | 22 | 9 | A48918 |
| 5 | 22 | 3.7 | 22 | 9 | AR047927 |
| 6 | 22 | 3.7 | 331801 | 3 | NMA42491 |
| 7 | 20 | 3.4 | 72056 | 62 | AC011183 |
| 8 | 20 | 3.4 | 91470 | 12 | AC011438 |

| | | | | | | |
|-----|----|-----|--------|----|-------------|-----------------------|
| 9 | 20 | 3.4 | 177816 | 65 | AC017643 | AC017643 Drosophila |
| 10 | 20 | 3.4 | 194897 | 4 | AC007697 | AC007697 Drosophila |
| 11 | 20 | 3.4 | 262395 | 5 | AE003802 | AE003802 Drosophila |
| 12 | 19 | 3.2 | 1152 | 14 | D89210 | D89210 Schistosoma |
| 13 | 19 | 3.2 | 1279 | 94 | AF273691 | AF273691 Mus musculus |
| 14 | 19 | 3.2 | 10128 | 1 | AE005021 | AE005021 Halobacter |
| 15 | 19 | 3.2 | 11873 | 9 | AR083116 | AR083116 Sequence |
| 16 | 19 | 3.2 | 11878 | 9 | AR083115 | AR083115 Sequence |
| 17 | 19 | 3.2 | 11883 | 9 | AR083114 | AR083114 Sequence |
| 18 | 19 | 3.2 | 36493 | 15 | SPBC262 | AL022103 S. pombe c |
| 19 | 19 | 3.2 | 82033 | 94 | AC068903 | AC068903 Mus musculus |
| 20 | 19 | 3.2 | 95292 | 94 | AC068952 | AC068952 Mus musculus |
| 21 | 19 | 3.2 | 110000 | 79 | AL554832_1 | Continuation (2 of |
| 22 | 19 | 3.2 | 111071 | 68 | AC024591 | AC024591 Homo sapi |
| 23 | 19 | 3.2 | 128562 | 66 | AC020962 | AC020962 Mus musculus |
| 24 | 19 | 3.2 | 130734 | 94 | AC073882 | AC073882 Mus musculus |
| 25 | 19 | 3.2 | 131888 | 93 | HSJ1057D4 | AL121777 Human DNA |
| 26 | 19 | 3.2 | 147009 | 76 | AC084179 | AC084179 Homo sapi |
| 27 | 19 | 3.2 | 16599 | 76 | AC083840 | AC083840 Homo sapi |
| 28 | 19 | 3.2 | 161549 | 64 | AC015815 | AC015815 Homo sapi |
| 29 | 19 | 3.2 | 170431 | 61 | AC009108 | AC009108 Homo sapi |
| 30 | 19 | 3.2 | 175249 | 63 | AC013343 | AC013343 Homo sapi |
| 31 | 19 | 3.2 | 179844 | 73 | AP002501 | AP002501 Homo sapi |
| 32 | 19 | 3.2 | 180717 | 73 | AC068169 | AC068169 Homo sapi |
| 33 | 19 | 3.2 | 181853 | 68 | AC024597 | AC024597 Homo sapi |
| 34 | 19 | 3.2 | 185892 | 78 | AC090652 | AC090652 Mus musculus |
| 35 | 19 | 3.2 | 191257 | 74 | AC069560 | AC069560 Mus musculus |
| 36 | 19 | 3.2 | 194575 | 81 | AL512653 | AL512653 Homo sapi |
| 37 | 19 | 3.2 | 198453 | 62 | AC011359 | AC011359 Homo sapi |
| 38 | 19 | 3.2 | 200860 | 81 | AL445211 | AL445211 Homo sapi |
| 39 | 19 | 3.2 | 235141 | 85 | AC004615 | AC004615 Homo sapi |
| 40 | 19 | 3.2 | 236050 | 79 | AL354813 | AL354813 Homo sapi |
| 41 | 18 | 3.0 | 18 | 9 | A48917 | A48917 Sequence 5 |
| 42 | 18 | 3.0 | 18 | 9 | A48917 | A48917 Sequence 8 |
| 43 | 18 | 3.0 | 18 | 9 | AR047923 | AR047923 Sequence |
| 44 | 18 | 3.0 | 18 | 9 | AR047926 | AR047926 Sequence |
| 45 | 18 | 3.0 | 1589 | 53 | AU049339 | AU049339 Rattus no |
| 46 | 18 | 3.0 | 1589 | 10 | AX086831 | AX086831 Sequence |
| 47 | 18 | 3.0 | 1589 | 93 | HSW801826 | AL168588 Homo sapi |
| 48 | 18 | 3.0 | 2520 | 9 | AR051480 | AR051480 Sequence |
| 49 | 18 | 3.0 | 2520 | 9 | AR072620 | AR072620 Sequence |
| 50 | 18 | 3.0 | 2520 | 9 | AR073165 | AR073165 Sequence |
| 51 | 18 | 3.0 | 2723 | 89 | AK026431 | AK026431 Homo sapi |
| 52 | 18 | 3.0 | 9450 | 59 | HPCCK09E1 | D63821 Hepatitis C |
| 53 | 18 | 3.0 | 10863 | 1 | AE005678 | AE005678 Calicivirus |
| 54 | 18 | 3.0 | 11031 | 1 | AE004648 | AE004648 Pseudomonas |
| 55 | 18 | 3.0 | 14704 | 1 | AE001970 | AE001970 Deinococcus |
| 56 | 18 | 3.0 | 38000 | 61 | AC010079 | AC010079 Leishmania |
| 57 | 18 | 3.0 | 42000 | 77 | AC087836 | AC087836 Leishmania |
| 58 | 18 | 3.0 | 58733 | 2 | AF217189 | AF217189 Sorangium |
| 59 | 18 | 3.0 | 68750 | 2 | AF210843 | AF210843 Sorangium |
| 60 | 18 | 3.0 | 80515 | 87 | AC011992 | AC011992 Homo sapi |
| 61 | 18 | 3.0 | 96424 | 14 | ATT22P11 | AL162971 Arabidopsis |
| 62 | 18 | 3.0 | 97749 | 85 | AC004906 | AC004906 Homo sapi |
| 63 | 18 | 3.0 | 110000 | 84 | LMFCHR32_03 | Continuation (4 of |
| 64 | 18 | 3.0 | 113345 | 88 | AF109075 | AF109075 Homo sapi |
| 65 | 18 | 3.0 | 115429 | 81 | AL500525 | AL500525 Homo sapi |
| 66 | 18 | 3.0 | 115925 | 75 | AC074128 | AC074128 Homo sapi |
| 67 | 18 | 3.0 | 135303 | 73 | AC068257 | AC068257 Homo sapi |
| 68 | 18 | 3.0 | 137506 | 85 | AC004824 | AC004824 Homo sapi |
| 69 | 18 | 3.0 | 144067 | 78 | AL138798 | AL138798 Homo sapi |
| 70 | 18 | 3.0 | 154461 | 69 | AC025711 | AC025711 Homo sapi |
| 71 | 18 | 3.0 | 154788 | 66 | AC021498 | AC021498 Homo sapi |
| 72 | 18 | 3.0 | 157349 | 78 | AF145206 | AF145206 Homo sapi |
| 73 | 18 | 3.0 | 157981 | 85 | AC005184 | AC005184 Homo sapi |
| 74 | 18 | 3.0 | 158427 | 79 | AL355516 | AL355516 Homo sapi |
| 75 | 18 | 3.0 | 162454 | 79 | AL158169 | AL158169 Homo sapi |
| 76 | 18 | 3.0 | 164679 | 86 | AC005921 | AC005921 Homo sapi |
| 77 | 18 | 3.0 | 169195 | 67 | AC023165 | AC023165 Homo sapi |
| 78 | 18 | 3.0 | 173081 | 78 | AL137001 | AL137001 Homo sapi |
| 79 | 18 | 3.0 | 177424 | 61 | AC009832 | AC009832 Homo sapi |
| 80 | 18 | 3.0 | 177720 | 69 | AC025680 | AC025680 Homo sapi |
| 81 | 18 | 3.0 | 180863 | 76 | AC079809 | AC079809 Homo sapi |
| 82 | 18 | 3.0 | 183235 | 66 | AC020714 | AC020714 Homo sapi |
| 83 | 18 | 3.0 | 207221 | 62 | AC012564 | AC012564 Homo sapi |
| 84 | 18 | 3.0 | 217730 | 77 | AC084823 | AC084823 Mus musculus |
| 85 | 18 | 3.0 | 218553 | 76 | AC079940 | AC079940 Mus musculus |
| 86 | 18 | 3.0 | 224376 | 61 | AC010538 | AC010538 Homo sapi |
| 87 | 18 | 3.0 | 326419 | 76 | AC079635 | AC079635 Mus musculus |
| 88 | 17 | 2.9 | 198 | 58 | AF065721 | AF065721 Stealith v |
| 89 | 17 | 2.9 | 339 | 93 | HSN052E17 | HSN052E17 Stealith v |
| 90 | 17 | 2.9 | 376 | 53 | AU026316 | AU026316 Rattus no |
| 91 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 92 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 93 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 94 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 95 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 96 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 97 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 98 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 99 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 100 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 101 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 102 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 103 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 104 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 105 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 106 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 107 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 108 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 109 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 110 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 111 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 112 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 113 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 114 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 115 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 116 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 117 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 118 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 119 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 120 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 121 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 122 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 123 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 124 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 125 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 126 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 127 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 128 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 129 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 130 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 131 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 132 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 133 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 134 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 135 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 136 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 137 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 138 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 139 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 140 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 141 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 142 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 143 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 144 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 145 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 146 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 147 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 148 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 149 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 150 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 151 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 152 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 153 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |
| 154 | 17 | 2.9 | 400 | 58 | AF174900 | AF174900 HIV-1 iso |

| | | | | | | | | | | | | | |
|-----|----|-----|--------|----|--------------|---------------------|-----|----|-----|--------|----|-----------|---------------------|
| 155 | 17 | 2.9 | 36849 | 3 | SC1C3 | AL023702 Streptomy | 228 | 17 | 2.9 | 155645 | 66 | AC021153 | AC021153 Homo sapi |
| 156 | 17 | 2.9 | 37056 | 85 | AC004700 | AC004700 Homo sapi | 229 | 17 | 2.9 | 155813 | 65 | AC019350 | AC019350 Homo sapi |
| 157 | 17 | 2.9 | 38024 | 6 | CEL52B10 | U41990 Caenorhabdi | 230 | 17 | 2.9 | 156241 | 80 | AL358034 | AL358034 Homo sapi |
| 158 | 17 | 2.9 | 39207 | 3 | SC4G1 | AL391039 Streptomy | 231 | 17 | 2.9 | 156557 | 63 | AC015545 | AC015545 Homo sapi |
| 159 | 17 | 2.9 | 39436 | 85 | AB000879 | AB000879 Homo sapi | 232 | 17 | 2.9 | 157189 | 4 | AC008307 | AC008307 Drosophill |
| 160 | 17 | 2.9 | 40121 | 3 | MSGY414A | AD000007 Mycobacte | 233 | 17 | 2.9 | 158447 | 75 | AC079071 | AC079071 Homo sapi |
| 161 | 17 | 2.9 | 40673 | 6 | LMFL377 | AL136326 Leishmani | 234 | 17 | 2.9 | 158446 | 86 | AC007395 | AC007395 Homo sapi |
| 162 | 17 | 2.9 | 40881 | 85 | AC004679 | AC004679 Homo sapi | 235 | 17 | 2.9 | 158645 | 79 | AL354708 | AL354708 Homo sapi |
| 163 | 17 | 2.9 | 42467 | 6 | CEL52B10 | U55373 Caenorhabdi | 236 | 17 | 2.9 | 159452 | 63 | AC015746 | AC015746 Homo sapi |
| 164 | 17 | 2.9 | 42584 | 89 | AF215835 | AF215839 Homo sapi | 237 | 17 | 2.9 | 160972 | 4 | AC007723 | AC007723 Drosophill |
| 165 | 17 | 2.9 | 43401 | 3 | MY13E12 | Z95390 Mycobacteri | 238 | 17 | 2.9 | 161087 | 65 | AC018850 | AC018850 Homo sapi |
| 166 | 17 | 2.9 | 43430 | 3 | MYC15B7 | Z81331 Mycobacteri | 239 | 17 | 2.9 | 162014 | 75 | AC074323 | AC074323 Homo sapi |
| 167 | 17 | 2.9 | 43600 | 85 | AC004209 | AC004209 Homo sapi | 240 | 17 | 2.9 | 162691 | 94 | AC003061 | AC003061 Mouse Chr |
| 168 | 17 | 2.9 | 44710 | 85 | AC004184 | AC004184 Homo sapi | 241 | 17 | 2.9 | 163853 | 13 | AP003285 | AP003285 Oryza sat |
| 169 | 17 | 2.9 | 45087 | 82 | AL590645 | AL590645 Homo sapi | 242 | 17 | 2.9 | 163570 | 87 | AC016910 | AC016910 Homo sapi |
| 170 | 17 | 2.9 | 46344 | 90 | AL356497 | AL356497 Human DNA | 243 | 17 | 2.9 | 163578 | 74 | AC073657 | AC073657 Homo sapi |
| 171 | 17 | 2.9 | 51082 | 78 | AC050454 | AC050454 Homo sapi | 244 | 17 | 2.9 | 164034 | 88 | AC073917 | AC073917 Homo sapi |
| 172 | 17 | 2.9 | 57399 | 91 | AP000433 | AP000433 Homo sapi | 245 | 17 | 2.9 | 164386 | 87 | AC011489 | AC011489 Homo sapi |
| 173 | 17 | 2.9 | 57597 | 77 | AC050035 | AC050035 Homo sapi | 246 | 17 | 2.9 | 164408 | 83 | CNS01DRF | AL117190 Homo sapi |
| 174 | 17 | 2.9 | 65174 | 69 | AC025214 | AC025214 Homo sapi | 247 | 17 | 2.9 | 165405 | 61 | AC009420 | AC009420 Homo sapi |
| 175 | 17 | 2.9 | 66066 | 70 | AC026566 | AC026566 Homo sapi | 248 | 17 | 2.9 | 165817 | 60 | AC007935 | AC007935 Homo sapi |
| 176 | 17 | 2.9 | 66686 | 77 | AC087757 | AC087757 Homo sapi | 249 | 17 | 2.9 | 166107 | 68 | AC024427 | AC024427 Homo sapi |
| 177 | 17 | 2.9 | 67392 | 65 | AC018284 | AC018284 Drosophill | 250 | 17 | 2.9 | 167237 | 66 | AC021609 | AC021609 Homo sapi |
| 178 | 17 | 2.9 | 74371 | 85 | AC005369 | AC005369 Homo sapi | 251 | 17 | 2.9 | 167310 | 73 | AC068972 | AC068972 Homo sapi |
| 179 | 17 | 2.9 | 80099 | 13 | AP002040 | AP002040 Arabidops | 252 | 17 | 2.9 | 167590 | 80 | AL356221 | AL356221 Homo sapi |
| 180 | 17 | 2.9 | 82604 | 91 | AL500522 | AL500522 Human DNA | 253 | 17 | 2.9 | 169241 | 72 | AL445987 | AL445987 Homo sapi |
| 181 | 17 | 2.9 | 84957 | 12 | AC006200 | AC006200 Arabidops | 254 | 17 | 2.9 | 169256 | 86 | AC048358 | AC048358 Homo sapi |
| 182 | 17 | 2.9 | 85657 | 12 | AC069473 | AC069473 Arabidops | 255 | 17 | 2.9 | 170030 | 86 | AC005829 | AC005829 Homo sapi |
| 183 | 17 | 2.9 | 86135 | 90 | AL159169 | AL159169 Human DNA | 256 | 17 | 2.9 | 170324 | 4 | AC007805 | AC007805 Drosophill |
| 184 | 17 | 2.9 | 88296 | 88 | AC027328 | AC027328 Homo sapi | 257 | 17 | 2.9 | 170457 | 76 | AC083944 | AC083944 Homo sapi |
| 185 | 17 | 2.9 | 88614 | 70 | AC026865 | AC026865 Homo sapi | 258 | 17 | 2.9 | 170711 | 71 | AC032020 | AC032020 Homo sapi |
| 186 | 17 | 2.9 | 88784 | 63 | AC012997 | AC012997 Drosophill | 259 | 17 | 2.9 | 170854 | 83 | AP002383 | AP002383 Homo sapi |
| 187 | 17 | 2.9 | 90767 | 84 | H0510A06 | AL442104 Oryza sat | 260 | 17 | 2.9 | 170856 | 73 | AC068160 | AC068160 Homo sapi |
| 188 | 17 | 2.9 | 94641 | 78 | AF257499 | AF257499 Homo sapi | 261 | 17 | 2.9 | 170891 | 61 | AL502118 | AL502118 Homo sapi |
| 189 | 17 | 2.9 | 100000 | 91 | AP000506 | AP000506 Homo sapi | 262 | 17 | 2.9 | 171539 | 80 | AL512788 | AL512788 Homo sapi |
| 190 | 17 | 2.9 | 100000 | 91 | AP000513 | AP000513 Homo sapi | 263 | 17 | 2.9 | 171775 | 70 | AC027015 | AC027015 Homo sapi |
| 191 | 17 | 2.9 | 107150 | 60 | AC008839 | AC008839 Homo sapi | 264 | 17 | 2.9 | 171790 | 62 | AC011106 | AC011106 Homo sapi |
| 192 | 17 | 2.9 | 107152 | 88 | AC020930 | AC020930 Homo sapi | 265 | 17 | 2.9 | 172004 | 86 | AC008482 | AC008482 Homo sapi |
| 193 | 17 | 2.9 | 109626 | 87 | AC008727 | AC008727 Homo sapi | 266 | 17 | 2.9 | 172312 | 80 | AL359882 | AL359882 Homo sapi |
| 194 | 17 | 2.9 | 109626 | 87 | AC010346 | AC010346 Homo sapi | 267 | 17 | 2.9 | 172650 | 62 | AC011042 | AC011042 Homo sapi |
| 195 | 17 | 2.9 | 110000 | 84 | LMFLCHR5_03 | Continuation (4 of | 268 | 17 | 2.9 | 173235 | 68 | AC023756 | AC023756 Homo sapi |
| 196 | 17 | 2.9 | 110000 | 84 | LMFLCHR5_03 | Continuation (4 of | 269 | 17 | 2.9 | 173254 | 68 | AC012057 | AC012057 Homo sapi |
| 197 | 17 | 2.9 | 110000 | 84 | LMFLCHR36_09 | Continuation (10 o | 270 | 17 | 2.9 | 173425 | 66 | AC021482 | AC021482 Homo sapi |
| 198 | 17 | 2.9 | 110000 | 84 | LMFLCHR36_11 | Continuation (12 o | 271 | 17 | 2.9 | 174517 | 77 | AC087831 | AC087831 Homo sapi |
| 199 | 17 | 2.9 | 110000 | 84 | LMFLCHR36_12 | Continuation (13 o | 272 | 17 | 2.9 | 174831 | 72 | AC060809 | AC060809 Homo sapi |
| 200 | 17 | 2.9 | 110603 | 93 | HS1421D16 | AL109940 Human DNA | 273 | 17 | 2.9 | 174852 | 61 | AC010195 | AC010195 Homo sapi |
| 201 | 17 | 2.9 | 113320 | 61 | AC010662 | AC010662 Drosophill | 274 | 17 | 2.9 | 176300 | 71 | AC027708 | AC027708 Homo sapi |
| 202 | 17 | 2.9 | 122258 | 80 | AL356427 | AL356427 Homo sapi | 275 | 17 | 2.9 | 177654 | 69 | AC025988 | AC025988 Homo sapi |
| 203 | 17 | 2.9 | 123085 | 90 | AL157406 | AL157406 Human DNA | 276 | 17 | 2.9 | 177656 | 76 | AC079911 | AC079911 Mus muscu |
| 204 | 17 | 2.9 | 126152 | 92 | HS433M19 | Z95703 Human DNA | 277 | 17 | 2.9 | 177826 | 64 | AC016852 | AC016852 Homo sapi |
| 205 | 17 | 2.9 | 132275 | 72 | AC048388 | AC048388 Homo sapi | 278 | 17 | 2.9 | 178395 | 81 | AL512664 | AL512664 Homo sapi |
| 206 | 17 | 2.9 | 133751 | 92 | HS1131C10 | AL121931 Human DNA | 279 | 17 | 2.9 | 178518 | 69 | AC025332 | AC025332 Homo sapi |
| 207 | 17 | 2.9 | 137928 | 78 | AC090419 | AC090419 Homo sapi | 280 | 17 | 2.9 | 179471 | 65 | AC078868 | AC078868 Homo sapi |
| 208 | 17 | 2.9 | 138141 | 65 | AC018433 | AC018433 Homo sapi | 281 | 17 | 2.9 | 182718 | 74 | AC017095 | AC017095 Homo sapi |
| 209 | 17 | 2.9 | 139505 | 93 | HSJ591C20 | AL118506 Human DNA | 282 | 17 | 2.9 | 183251 | 63 | AC013555 | AC013555 Homo sapi |
| 210 | 17 | 2.9 | 143759 | 88 | AC073532 | AC073532 Homo sapi | 283 | 17 | 2.9 | 184165 | 60 | AC008763 | AC008763 Homo sapi |
| 211 | 17 | 2.9 | 145642 | 71 | AC027805 | AC027805 Homo sapi | 284 | 17 | 2.9 | 184651 | 81 | AL450317 | AL450317 Mus muscu |
| 212 | 17 | 2.9 | 146018 | 65 | AC018404 | AC018404 Homo sapi | 285 | 17 | 2.9 | 185260 | 64 | AC016065 | AC016065 Homo sapi |
| 213 | 17 | 2.9 | 146393 | 70 | AC026791 | AC026791 Homo sapi | 286 | 17 | 2.9 | 185551 | 93 | HSJ171K16 | AL121881 Human DNA |
| 214 | 17 | 2.9 | 148960 | 85 | AB000882 | AB000882 Homo sapi | 287 | 17 | 2.9 | 185948 | 87 | AC022370 | AC022370 Homo sapi |
| 215 | 17 | 2.9 | 149143 | 62 | AC011591 | AC011591 Homo sapi | 288 | 17 | 2.9 | 186102 | 83 | AP001983 | AP001983 Homo sapi |
| 216 | 17 | 2.9 | 150250 | 70 | AC027573 | AC027573 Homo sapi | 289 | 17 | 2.9 | 187965 | 64 | AC017010 | AC017010 Homo sapi |
| 217 | 17 | 2.9 | 151289 | 83 | AP003228 | AP003228 Oryza sat | 290 | 17 | 2.9 | 188463 | 74 | AC073413 | AC073413 Homo sapi |
| 218 | 17 | 2.9 | 151496 | 83 | AP003312 | AP003312 Oryza sat | 291 | 17 | 2.9 | 188526 | 62 | AC012342 | AC012342 Homo sapi |
| 219 | 17 | 2.9 | 151497 | 71 | AC027809 | AC027809 Homo sapi | 292 | 17 | 2.9 | 188713 | 88 | AC024167 | AC024167 Homo sapi |
| 220 | 17 | 2.9 | 152159 | 70 | AC027334 | AC027334 Homo sapi | 293 | 17 | 2.9 | 188755 | 87 | AC013562 | AC013562 Homo sapi |
| 221 | 17 | 2.9 | 152513 | 83 | AP003207 | AP003207 Oryza sat | 294 | 17 | 2.9 | 188812 | 63 | AC013419 | AC013419 Homo sapi |
| 222 | 17 | 2.9 | 152517 | 87 | AC009468 | AC009468 Homo sapi | 295 | 17 | 2.9 | 189333 | 76 | AC079445 | AC079445 Mus muscu |
| 223 | 17 | 2.9 | 154685 | 86 | AC007214 | AC007214 Pan trogl | 296 | 17 | 2.9 | 189786 | 73 | AC068015 | AC068015 Homo sapi |
| 224 | 17 | 2.9 | 154745 | 61 | AC010860 | AC010860 Homo sapi | 297 | 17 | 2.9 | 190386 | 75 | AC074216 | AC074216 Homo sapi |
| 225 | 17 | 2.9 | 155068 | 70 | AC027307 | AC027307 Homo sapi | 298 | 17 | 2.9 | 190844 | 75 | AC073990 | AC073990 Homo sapi |
| 226 | 17 | 2.9 | 155258 | 84 | HS66C13 | AL354047 Homo sapi | 299 | 17 | 2.9 | 191111 | 86 | AC007458 | AC007458 Homo sapi |
| 227 | 17 | 2.9 | 155399 | 69 | AC025663 | AC025663 Homo sapi | 300 | 17 | 2.9 | 192431 | 67 | AC022059 | AC022059 Homo sapi |

| | | | | | | | | | | | | |
|-------|----|-----|--------|----|------------|-----|----|-----|-----|----|------------|--------------------|
| C 301 | 17 | 2.9 | 192868 | 88 | AC023484 | 374 | 16 | 2.7 | 343 | 59 | HIV1S31756 | L13574 Human Immun |
| C 302 | 17 | 2.9 | 193098 | 89 | AL137140 | 375 | 16 | 2.7 | 343 | 59 | HIV1S31757 | L13575 Human Immun |
| C 303 | 17 | 2.9 | 193426 | 73 | AC067942 | 376 | 16 | 2.7 | 343 | 59 | HIV1S31758 | L13525 Human Immun |
| C 304 | 17 | 2.9 | 193505 | 79 | AL139818 | 377 | 16 | 2.7 | 343 | 59 | HIV1S37R10 | L13534 Human Immun |
| C 305 | 17 | 2.9 | 193816 | 69 | AC025874 | 378 | 16 | 2.7 | 343 | 59 | HIV1S37R11 | L13533 Human Immun |
| C 306 | 17 | 2.9 | 194151 | 66 | AC021185 | 379 | 16 | 2.7 | 343 | 59 | HIV1S37R12 | L13536 Human Immun |
| C 307 | 17 | 2.9 | 195310 | 68 | AC024189 | 380 | 16 | 2.7 | 343 | 59 | HIV1S37R2 | L13526 Human Immun |
| C 308 | 17 | 2.9 | 195466 | 71 | AC032024 | 381 | 16 | 2.7 | 343 | 59 | HIV1S37R3 | L13527 Human Immun |
| C 309 | 17 | 2.9 | 195472 | 65 | AC018398 | 382 | 16 | 2.7 | 343 | 59 | HIV1S37R4 | L13528 Human Immun |
| C 310 | 17 | 2.9 | 195556 | 65 | AC018382 | 383 | 16 | 2.7 | 343 | 59 | HIV1S37R5 | L13529 Human Immun |
| C 311 | 17 | 2.9 | 197871 | 73 | AC068053 | 384 | 16 | 2.7 | 343 | 59 | HIV1S37R6 | L13530 Human Immun |
| C 312 | 17 | 2.9 | 198388 | 74 | AC073426 | 385 | 16 | 2.7 | 343 | 59 | HIV1S37R7 | L13531 Human Immun |
| C 313 | 17 | 2.9 | 198677 | 1 | AE001863 | 386 | 16 | 2.7 | 343 | 59 | HIV1S37R8 | L13532 Human Immun |
| C 314 | 17 | 2.9 | 199314 | 61 | AC009097 | 387 | 16 | 2.7 | 343 | 59 | HIV1S37R9 | L13533 Human Immun |
| C 315 | 17 | 2.9 | 199853 | 75 | AC073785 | 388 | 16 | 2.7 | 343 | 59 | HIV1B05G1 | L36011 Human Immun |
| C 316 | 17 | 2.9 | 200033 | 93 | HSa251973 | 389 | 16 | 2.7 | 343 | 59 | HIV1B05G2 | L36012 Human Immun |
| C 317 | 17 | 2.9 | 200033 | 69 | AC025105 | 390 | 16 | 2.7 | 344 | 58 | AF028881 | AF028881 HIV-1 L50 |
| C 318 | 17 | 2.9 | 201846 | 60 | AC009036 | 391 | 16 | 2.7 | 354 | 58 | AF224030 | AF224030 HIV-1 C25 |
| C 319 | 17 | 2.9 | 207083 | 72 | AC055797 | 392 | 16 | 2.7 | 354 | 58 | AF020207 | AF020207 Rattus no |
| C 320 | 17 | 2.9 | 209034 | 75 | AC073789 | 393 | 16 | 2.7 | 356 | 58 | AF050932 | AF050932 HIV-1 L50 |
| C 321 | 17 | 2.9 | 212209 | 68 | AC024651 | 394 | 16 | 2.7 | 357 | 58 | AF191457 | AF191457 HIV-1 L50 |
| C 322 | 17 | 2.9 | 217115 | 60 | AC007305 | 395 | 16 | 2.7 | 360 | 58 | AF073390 | AF073390 HIV-1 L50 |
| C 323 | 17 | 2.9 | 220995 | 61 | AC010687 | 396 | 16 | 2.7 | 360 | 58 | AF224046 | AF224046 HIV-1 C26 |
| C 324 | 17 | 2.9 | 221913 | 67 | AC022129 | 397 | 16 | 2.7 | 368 | 58 | AF017866 | AF017866 HIV-1 Pat |
| C 325 | 17 | 2.9 | 223000 | 6 | LMFCHRA4 | 398 | 16 | 2.7 | 372 | 58 | AF050938 | AF050938 HIV-1 L50 |
| C 326 | 17 | 2.9 | 226483 | 5 | AE003693 | 399 | 16 | 2.7 | 372 | 58 | AF020208 | AF020208 Rattus no |
| C 327 | 17 | 2.9 | 228859 | 66 | AC021141 | 400 | 16 | 2.7 | 375 | 58 | AF050937 | AF050937 HIV-1 L50 |
| C 328 | 17 | 2.9 | 229122 | 5 | AE003700 | 401 | 16 | 2.7 | 375 | 58 | AF050939 | AF050939 HIV-1 L50 |
| C 329 | 17 | 2.9 | 229252 | 81 | AL390732 | 402 | 16 | 2.7 | 380 | 58 | AF050935 | AF050935 HIV-1 L50 |
| C 330 | 17 | 2.9 | 230651 | 74 | AC073535 | 403 | 16 | 2.7 | 385 | 58 | AF073402 | AF073402 HIV-1 L50 |
| C 331 | 17 | 2.9 | 233752 | 74 | AC073583 | 404 | 16 | 2.7 | 386 | 58 | AF028884 | AF028884 HIV-1 L50 |
| C 332 | 17 | 2.9 | 240285 | 60 | AC008939 | 405 | 16 | 2.7 | 386 | 58 | AF028887 | AF028887 HIV-1 L50 |
| C 333 | 17 | 2.9 | 240825 | 10 | AX087869 | 406 | 16 | 2.7 | 388 | 58 | AF014238 | AF014238 HIV-1 Pat |
| C 334 | 17 | 2.9 | 258829 | 75 | AC073771 | 407 | 16 | 2.7 | 388 | 58 | AF014239 | AF014239 HIV-1 Pat |
| C 335 | 17 | 2.9 | 259421 | 74 | AC073689 | 408 | 16 | 2.7 | 388 | 58 | AF014240 | AF014240 HIV-1 Pat |
| C 336 | 17 | 2.9 | 267578 | 86 | AC006449 | 409 | 16 | 2.7 | 388 | 58 | AF014344 | AF014344 HIV-1 Pat |
| C 337 | 17 | 2.9 | 291136 | 2 | AP002556 | 410 | 16 | 2.7 | 388 | 58 | AF014345 | AF014345 HIV-1 Pat |
| C 338 | 17 | 2.9 | 294308 | 4 | AE003595 | 411 | 16 | 2.7 | 388 | 58 | AF014358 | AF014358 HIV-1 Pat |
| C 339 | 17 | 2.9 | 303250 | 2 | AP001518 | 412 | 16 | 2.7 | 390 | 58 | AF015406 | AF015406 HIV-1 Pat |
| C 340 | 17 | 2.9 | 310360 | 75 | AC074172 | 413 | 16 | 2.7 | 390 | 58 | AF062050 | AF062050 HIV-1 L50 |
| C 341 | 17 | 2.9 | 310565 | 83 | CEX62P5 | 414 | 16 | 2.7 | 391 | 58 | AF014194 | AF014194 HIV-1 Pat |
| C 342 | 17 | 2.9 | 338534 | 2 | ECOMW93 | 415 | 16 | 2.7 | 391 | 58 | AF014199 | AF014199 HIV-1 Pat |
| C 343 | 17 | 2.9 | 339485 | 13 | AP172282 | 416 | 16 | 2.7 | 391 | 58 | AF014202 | AF014202 HIV-1 Pat |
| C 344 | 17 | 2.9 | 340000 | 91 | AP001675 | 417 | 16 | 2.7 | 391 | 58 | AF014216 | AF014216 HIV-1 Pat |
| C 345 | 16 | 2.7 | 41 | 10 | E08211 | 418 | 16 | 2.7 | 391 | 58 | AF014225 | AF014225 HIV-1 Pat |
| C 346 | 16 | 2.7 | 222 | 10 | E08211 | 419 | 16 | 2.7 | 391 | 58 | AF014228 | AF014228 HIV-1 Pat |
| C 347 | 16 | 2.7 | 222 | 14 | OSTOSI1A | 420 | 16 | 2.7 | 391 | 58 | AF014233 | AF014233 HIV-1 Pat |
| C 348 | 16 | 2.7 | 224 | 14 | ROALPTS2 | 421 | 16 | 2.7 | 391 | 58 | AF050899 | AF050899 HIV-1 mat |
| C 349 | 16 | 2.7 | 224 | 14 | ROALPTS2 | 422 | 16 | 2.7 | 391 | 58 | AF050900 | AF050900 HIV-1 mat |
| C 350 | 16 | 2.7 | 224 | 14 | ROBRATTS2 | 423 | 16 | 2.7 | 391 | 58 | AF050927 | AF050927 HIV-1 L50 |
| C 351 | 16 | 2.7 | 224 | 14 | ROBRATTS2 | 424 | 16 | 2.7 | 391 | 58 | AF050928 | AF050928 HIV-1 L50 |
| C 352 | 16 | 2.7 | 224 | 14 | ROBRATTS2 | 425 | 16 | 2.7 | 391 | 58 | AF050929 | AF050929 HIV-1 L50 |
| C 353 | 16 | 2.7 | 224 | 14 | ROBRATTS2 | 426 | 16 | 2.7 | 391 | 58 | AF050931 | AF050931 HIV-1 L50 |
| C 354 | 16 | 2.7 | 225 | 14 | CGRACITS2 | 427 | 16 | 2.7 | 391 | 58 | AF050936 | AF050936 HIV-1 L50 |
| C 355 | 16 | 2.7 | 225 | 14 | CGRACITS2 | 428 | 16 | 2.7 | 391 | 58 | AF050942 | AF050942 HIV-1 L50 |
| C 356 | 16 | 2.7 | 225 | 14 | ROBRATTS2 | 429 | 16 | 2.7 | 391 | 58 | AF050944 | AF050944 HIV-1 L50 |
| C 357 | 16 | 2.7 | 225 | 14 | ROBRATTS2 | 430 | 16 | 2.7 | 391 | 58 | AF050946 | AF050946 HIV-1 L50 |
| C 358 | 16 | 2.7 | 225 | 14 | ROBRATTS2 | 431 | 16 | 2.7 | 391 | 58 | AF050951 | AF050951 HIV-1 L50 |
| C 359 | 16 | 2.7 | 225 | 14 | ROBRATTS2 | 432 | 16 | 2.7 | 391 | 58 | AF050952 | AF050952 HIV-1 L50 |
| C 360 | 16 | 2.7 | 225 | 14 | ROBRATTS2 | 433 | 16 | 2.7 | 391 | 58 | AF050953 | AF050953 HIV-1 L50 |
| C 361 | 16 | 2.7 | 225 | 14 | ROBRATTS2 | 434 | 16 | 2.7 | 391 | 58 | AF050959 | AF050959 HIV-1 L50 |
| C 362 | 16 | 2.7 | 225 | 14 | ROBRATTS2 | 435 | 16 | 2.7 | 391 | 58 | AF050960 | AF050960 HIV-1 L50 |
| C 363 | 16 | 2.7 | 225 | 14 | ROBRATTS2 | 436 | 16 | 2.7 | 391 | 58 | AF051008 | AF051008 HIV-1 L50 |
| C 364 | 16 | 2.7 | 293 | 3 | G04984 | 437 | 16 | 2.7 | 391 | 58 | AF051009 | AF051009 HIV-1 L50 |
| C 365 | 16 | 2.7 | 334 | 9 | AX000736 | 438 | 16 | 2.7 | 391 | 58 | AF051010 | AF051010 HIV-1 L50 |
| C 366 | 16 | 2.7 | 336 | 59 | HIV1E7G4 | 439 | 16 | 2.7 | 391 | 58 | AF051012 | AF051012 HIV-1 L50 |
| C 367 | 16 | 2.7 | 341 | 58 | AF028579 | 440 | 16 | 2.7 | 391 | 58 | AF051013 | AF051013 HIV-1 L50 |
| C 368 | 16 | 2.7 | 343 | 58 | AF028578 | 441 | 16 | 2.7 | 391 | 58 | AF051014 | AF051014 HIV-1 L50 |
| C 369 | 16 | 2.7 | 343 | 59 | HIV1S317S1 | 442 | 16 | 2.7 | 391 | 58 | AF051016 | AF051016 HIV-1 L50 |
| C 370 | 16 | 2.7 | 343 | 59 | HIV1S317S2 | 443 | 16 | 2.7 | 391 | 58 | AF051019 | AF051019 HIV-1 L50 |
| C 371 | 16 | 2.7 | 343 | 59 | HIV1S317S3 | 444 | 16 | 2.7 | 391 | 58 | AF051022 | AF051022 HIV-1 L50 |
| C 372 | 16 | 2.7 | 343 | 59 | HIV1S317S4 | 445 | 16 | 2.7 | 391 | 58 | AF051024 | AF051024 HIV-1 L50 |
| C 373 | 16 | 2.7 | 343 | 59 | HIV1S317S5 | 446 | 16 | 2.7 | 391 | 58 | AF051027 | AF051027 HIV-1 L50 |

| | | | | | | | | | | | | | |
|-----|----|-----|-----|-----|------------|--------------------|-----|----|-----|------|----|-----------|-----------------------|
| 447 | 16 | 2.7 | 391 | 58 | AF051030 | HIV-1 iso | 520 | 16 | 2.7 | 626 | 13 | AF202406 | AF202406 Roscoea h |
| 448 | 16 | 2.7 | 393 | 94 | AF020209 | Rattus no | 521 | 16 | 2.7 | 627 | 13 | AF202376 | AF202376 Hedychium |
| 449 | 16 | 2.7 | 394 | 58 | AF014269 | HIV-1 pat | 522 | 16 | 2.7 | 627 | 13 | AF202380 | AF202380 Hedychium |
| 450 | 16 | 2.7 | 396 | 58 | AF084766 | HIV-1 pat | 523 | 16 | 2.7 | 627 | 13 | AF202381 | AF202381 Hedychium |
| 451 | 16 | 2.7 | 396 | 58 | AF084767 | HIV-1 pat | 524 | 16 | 2.7 | 627 | 13 | AF202382 | AF202382 Hedychium |
| 452 | 16 | 2.7 | 396 | 58 | AF084768 | HIV-1 pat | 525 | 16 | 2.7 | 627 | 13 | AF202393 | AF202393 Hedychium |
| 453 | 16 | 2.7 | 396 | 58 | AF084769 | HIV-1 pat | 526 | 16 | 2.7 | 627 | 13 | AF202394 | AF202394 Hedychium |
| 454 | 16 | 2.7 | 396 | 58 | AF084770 | HIV-1 pat | 527 | 16 | 2.7 | 627 | 13 | AF202397 | AF202397 Hedychium |
| 455 | 16 | 2.7 | 396 | 58 | AF084771 | HIV-1 pat | 528 | 16 | 2.7 | 628 | 13 | AF202377 | AF202377 Hedychium |
| 456 | 16 | 2.7 | 396 | 58 | AF084772 | HIV-1 pat | 529 | 16 | 2.7 | 628 | 13 | AF202378 | AF202378 Hedychium |
| 457 | 16 | 2.7 | 396 | 58 | AF084773 | HIV-1 pat | 530 | 16 | 2.7 | 628 | 13 | AF202383 | AF202383 Hedychium |
| 458 | 16 | 2.7 | 396 | 58 | AF084774 | HIV-1 pat | 531 | 16 | 2.7 | 628 | 13 | AF202384 | AF202384 Hedychium |
| 459 | 16 | 2.7 | 396 | 58 | AF084775 | HIV-1 pat | 532 | 16 | 2.7 | 628 | 13 | AF202389 | AF202389 Hedychium |
| 460 | 16 | 2.7 | 396 | 58 | AF084776 | HIV-1 pat | 533 | 16 | 2.7 | 628 | 13 | AF202390 | AF202390 Hedychium |
| 461 | 16 | 2.7 | 396 | 58 | AF084777 | HIV-1 pat | 534 | 16 | 2.7 | 628 | 13 | AF202391 | AF202391 Hedychium |
| 462 | 16 | 2.7 | 396 | 59 | HIV1DUMC | | 535 | 16 | 2.7 | 628 | 13 | AF202392 | AF202392 Hedychium |
| 463 | 16 | 2.7 | 396 | 59 | HIV1DUMK | | 536 | 16 | 2.7 | 628 | 13 | AF202396 | AF202396 Hedychium |
| 464 | 16 | 2.7 | 397 | 58 | AF014195 | HIV-1 iso | 537 | 16 | 2.7 | 629 | 13 | AF202379 | AF202379 Hedychium |
| 465 | 16 | 2.7 | 399 | 58 | AF174897 | HIV-1 iso | 538 | 16 | 2.7 | 629 | 13 | AF202402 | AF202402 Hedychium |
| 466 | 16 | 2.7 | 399 | 58 | AF174898 | HIV-1 iso | 539 | 16 | 2.7 | 631 | 13 | AF202401 | AF202401 Hedychium |
| 467 | 16 | 2.7 | 399 | 58 | AF174899 | HIV-1 iso | 540 | 16 | 2.7 | 651 | 97 | HUM0288BF | HUM0288BF Human CPG 1 |
| 468 | 16 | 2.7 | 400 | 58 | AF014222 | HIV-1 pat | 541 | 16 | 2.7 | 729 | 58 | AF071291 | AF071291 HIV-1 iso |
| 469 | 16 | 2.7 | 400 | 58 | AF174851 | HIV-1 iso | 542 | 16 | 2.7 | 729 | 58 | AF071300 | AF071300 HIV-1 iso |
| 470 | 16 | 2.7 | 400 | 58 | AF174895 | HIV-1 iso | 543 | 16 | 2.7 | 729 | 58 | AF196682 | AF196682 HIV-1 iso |
| 471 | 16 | 2.7 | 400 | 58 | AF174896 | HIV-1 iso | 544 | 16 | 2.7 | 738 | 58 | AF071302 | AF071302 HIV-1 iso |
| 472 | 16 | 2.7 | 400 | 58 | AF174902 | HIV-1 iso | 545 | 16 | 2.7 | 738 | 58 | AF071308 | AF071308 HIV-1 iso |
| 473 | 16 | 2.7 | 400 | 58 | AF174962 | HIV-1 iso | 546 | 16 | 2.7 | 759 | 96 | TVA250739 | TVA250739 Trypanoso |
| 474 | 16 | 2.7 | 400 | 58 | AF174964 | HIV-1 iso | 547 | 16 | 2.7 | 774 | 9 | A70164 | A70164 Sequence 25 |
| 475 | 16 | 2.7 | 400 | 58 | AF174965 | HIV-1 iso | 548 | 16 | 2.7 | 774 | 14 | MAZ93103 | MAZ93103 M.acuminata |
| 476 | 16 | 2.7 | 400 | 58 | AF174966 | HIV-1 iso | 549 | 16 | 2.7 | 790 | 53 | CNS07270 | AF142866 clone BA0 |
| 477 | 16 | 2.7 | 400 | 58 | AF174967 | HIV-1 iso | 550 | 16 | 2.7 | 843 | 93 | CNS072W | AF142676 clone BA0 |
| 478 | 16 | 2.7 | 400 | 58 | AF174968 | HIV-1 iso | 551 | 16 | 2.7 | 863 | 91 | BC005052 | BC005052 Homo sapi |
| 479 | 16 | 2.7 | 400 | 58 | AF174969 | HIV-1 iso | 552 | 16 | 2.7 | 871 | 97 | HUM17SAG | ME07542 Human histo |
| 480 | 16 | 2.7 | 400 | 58 | AF174970 | HIV-1 iso | 553 | 16 | 2.7 | 886 | 53 | CNS0750E | AL430428 clone BA0 |
| 481 | 16 | 2.7 | 400 | 58 | AF174971 | HIV-1 iso | 554 | 16 | 2.7 | 892 | 72 | AC052426 | AC052426 clone BA0 |
| 482 | 16 | 2.7 | 401 | 58 | AF028583 | HIV-1 iso | 555 | 16 | 2.7 | 910 | 75 | AC075071 | AC075071 Giardia 1 |
| 483 | 16 | 2.7 | 401 | 58 | AF073401 | HIV-1 iso | 556 | 16 | 2.7 | 926 | 53 | CNS0730B | AL426897 clone BA0 |
| 484 | 16 | 2.7 | 406 | 58 | AF073400 | HIV-1 iso | 557 | 16 | 2.7 | 952 | 76 | AC083683 | AC083683 Giardia 1 |
| 485 | 16 | 2.7 | 409 | 58 | AF014214 | HIV-1 pat | 558 | 16 | 2.7 | 966 | 53 | CNS07630 | AL430906 clone XBA |
| 486 | 16 | 2.7 | 411 | 94 | AF020210 | Rattus no | 559 | 16 | 2.7 | 990 | 10 | AX089467 | AX089467 Sequence |
| 487 | 16 | 2.7 | 412 | 58 | AF014335 | HIV-1 pat | 560 | 16 | 2.7 | 999 | 74 | AC070320 | AC070320 Giardia 1 |
| 488 | 16 | 2.7 | 414 | 4 | AB040074 | Hydra mag | 561 | 16 | 2.7 | 1022 | 76 | AC080628 | AC080628 Giardia 1 |
| 489 | 16 | 2.7 | 414 | 58 | AF028577 | HIV-1 iso | 562 | 16 | 2.7 | 1040 | 72 | AC060129 | AC060129 Giardia 1 |
| 490 | 16 | 2.7 | 420 | 58 | AF028580 | HIV-1 iso | 563 | 16 | 2.7 | 1061 | 53 | CNS0738M | AL427196 clone BA0 |
| 491 | 16 | 2.7 | 448 | 58 | AF073403 | HIV-1 iso | 564 | 16 | 2.7 | 1150 | 7 | OC077896 | U77896 Oryctolagus |
| 492 | 16 | 2.7 | 454 | 58 | AF028586 | HIV-1 iso | 565 | 16 | 2.7 | 1247 | 85 | AB051152 | AB051152 Macaca fa |
| 493 | 16 | 2.7 | 480 | 14 | CNS01BXR | Botrytis | 566 | 16 | 2.7 | 1292 | 3 | UB085123 | UB8123 Unidentifite |
| 494 | 16 | 2.7 | 492 | 3 | SCSCYGEN | X97833 S.cyanogenu | 567 | 16 | 2.7 | 1299 | 5 | AF017268 | AF017268 Toxoplas |
| 495 | 16 | 2.7 | 492 | 3 | SGSGHE | X97837 S.glaucense | 568 | 16 | 2.7 | 1314 | 9 | AF072866 | AR072866 Sequence |
| 496 | 16 | 2.7 | 497 | 504 | 97 | SGSGHE | 569 | 16 | 2.7 | 1314 | 10 | E14955 | E14955 Aspergillus |
| 497 | 16 | 2.7 | 504 | 97 | HUMQ086H04 | AF075094 Homo sapi | 570 | 16 | 2.7 | 1314 | 10 | E14955 | E14955 Aspergillus |
| 498 | 16 | 2.7 | 523 | 9 | A93100 | Sequence 4 | 571 | 16 | 2.7 | 1314 | 10 | E27570 | E27570 Thermotabl |
| 499 | 16 | 2.7 | 559 | 14 | RP0388304 | AJ31080 Roscoea p | 572 | 16 | 2.7 | 1314 | 13 | ATFA0A | Y09040 A.terreus m |
| 500 | 16 | 2.7 | 561 | 14 | CSP38301 | Caulileya | 573 | 16 | 2.7 | 1381 | 94 | AF273690 | AF273690 Mus muscu |
| 501 | 16 | 2.7 | 562 | 14 | HMU388297 | AJ388297 Hedychium | 574 | 16 | 2.7 | 1417 | 96 | TG064517 | U64517 Toxoplasma |
| 502 | 16 | 2.7 | 563 | 14 | CGR388299 | AJ388299 Caulileya | 575 | 16 | 2.7 | 1449 | 89 | AF287302 | AF287302 Homo sapi |
| 503 | 16 | 2.7 | 563 | 14 | CGR388299 | AJ388299 Hedychium | 576 | 16 | 2.7 | 1455 | 97 | IR1845189 | AL350507 Homo sapi |
| 504 | 16 | 2.7 | 574 | 54 | HS421120S | AL035140 H.sapiens | 577 | 16 | 2.7 | 1468 | 89 | AF119879 | X91008 T.thermophi |
| 505 | 16 | 2.7 | 602 | 5 | AF320102 | Dendronep | 578 | 16 | 2.7 | 1477 | 59 | HIV1205 | AF119879 Homo sapi |
| 506 | 16 | 2.7 | 603 | 5 | AF320102 | Dendronep | 579 | 16 | 2.7 | 1513 | 10 | AX082314 | L11785 Human immun |
| 507 | 16 | 2.7 | 603 | 5 | AF320103 | Dendronep | 580 | 16 | 2.7 | 1519 | 94 | AF273689 | AX082314 Sequence |
| 508 | 16 | 2.7 | 604 | 5 | AF320098 | Dendronep | 581 | 16 | 2.7 | 1525 | 94 | AB005049 | AF273689 Mus muscu |
| 509 | 16 | 2.7 | 604 | 5 | AF320101 | Dendronep | 582 | 16 | 2.7 | 1581 | 94 | AF134488 | AB005049 Cavia por |
| 510 | 16 | 2.7 | 608 | 54 | G39591 | Zebira | 583 | 16 | 2.7 | 1600 | 3 | RCAN1FQ | AF134488 Cavia por |
| 511 | 16 | 2.7 | 610 | 5 | AF320099 | Dendronep | 584 | 16 | 2.7 | 1669 | 94 | AF273686 | M26323 Rhododactyl |
| 512 | 16 | 2.7 | 612 | 13 | AF202374 | Hedychium | 585 | 16 | 2.7 | 1669 | 94 | AF273687 | AF273686 Mus muscu |
| 513 | 16 | 2.7 | 621 | 13 | AF202403 | Hedychium | 586 | 16 | 2.7 | 1699 | 94 | AF273687 | AF273687 Mus muscu |
| 514 | 16 | 2.7 | 621 | 13 | AF202403 | Hedychium | 587 | 16 | 2.7 | 1723 | 94 | AF273688 | X13597 Streptomyce |
| 515 | 16 | 2.7 | 623 | 13 | AF202405 | Pommeresc | 588 | 16 | 2.7 | 1723 | 94 | AF273688 | AF273688 Mus muscu |
| 516 | 16 | 2.7 | 623 | 13 | AF202419 | Caulileya | 589 | 16 | 2.7 | 1788 | 92 | HSBS12 | X12433 Human pHS1- |
| 517 | 16 | 2.7 | 626 | 13 | AF202385 | Hedychium | 590 | 16 | 2.7 | 1790 | 12 | AF118126 | AF118126 Megnaport |
| 518 | 16 | 2.7 | 626 | 13 | AF202388 | Hedychium | 591 | 16 | 2.7 | 1803 | 10 | AX090436 | AX090436 Sequence |
| 519 | 16 | 2.7 | 626 | 13 | AF202395 | Hedychium | 592 | 16 | 2.7 | 1839 | 12 | AB045892 | AB045892 Neperthes |

| | | | | | | | | | | | |
|-----|----|-----|-------|----|------------|---------------------|-----|--------|----|-----------|----------------------|
| 739 | 16 | 2.7 | 17063 | 85 | AB031007 | AB031007 Homo sapi | 812 | 52797 | 11 | AF068845 | AF068845 Mycobacte |
| 740 | 16 | 2.7 | 17286 | 91 | D49493 | D49493 Human gene | 813 | 55700 | 65 | AC017867 | AC017867 Drosophill |
| 741 | 16 | 2.7 | 18221 | 64 | AC017207 | AC017207 Drosophill | 814 | 56999 | 60 | AC008776 | AC008776 Homo sapi |
| 742 | 16 | 2.7 | 18510 | 65 | AC017804 | AC017804 Drosophill | 815 | 57214 | 65 | AC019870 | AC019870 Homo sapi |
| 743 | 16 | 2.7 | 20413 | 13 | AF271736S1 | AF271736 Englena g | 816 | 57306 | 4 | AC005416 | AC005416 Drosophill |
| 744 | 16 | 2.7 | 20413 | 85 | AC004631 | AC004631 Homo sapi | 817 | 57906 | 14 | NCB7319 | AL389899 Neurospor |
| 745 | 16 | 2.7 | 20823 | 85 | CELT24H7 | U28940 Caenorhabdi | 818 | 58220 | 78 | AC091188 | AC091188 Homo sapi |
| 746 | 16 | 2.7 | 21040 | 13 | AF271738S2 | AF271739 Englena g | 819 | 59264 | 65 | AC018278 | AC018278 Homo sapi |
| 747 | 16 | 2.7 | 21721 | 9 | AX000057 | AX000057 Sequence | 820 | 59747 | 89 | AF119117 | AF119117 Homo sapi |
| 748 | 16 | 2.7 | 22789 | 3 | SC2A11 | AL031184 Streptomy | 821 | 60040 | 4 | AC005128 | AC005128 Drosophill |
| 749 | 16 | 2.7 | 22898 | 6 | CELF35A5 | U46675 Caenorhabdi | 822 | 60966 | 85 | AC003030 | AC003030 Homo sapi |
| 750 | 16 | 2.7 | 22976 | 9 | AX000035 | AX000035 Sequence | 823 | 61133 | 86 | AC005838 | AC005838 Homo sapi |
| 751 | 16 | 2.7 | 23187 | 10 | AX1000598 | AX1000598 Sequence | 824 | 61845 | 78 | AC090837 | AC090837 Homo sapi |
| 752 | 16 | 2.7 | 23222 | 3 | SCGH18 | AL357152 Streptomy | 825 | 61969 | 63 | AC014103 | AC014103 Drosophill |
| 753 | 16 | 2.7 | 23952 | 6 | CEP29D10 | Z75952 Caenorhabdi | 826 | 63397 | 90 | AL353604 | AL353604 Human DNA |
| 754 | 16 | 2.7 | 24700 | 3 | SC68 | AL033654 Streptomy | 827 | 64120 | 63 | AC014287 | AC014287 Drosophill |
| 755 | 16 | 2.7 | 25359 | 6 | CEFO2D10 | AL033654 Streptomy | 828 | 65318 | 77 | AC084339 | AC084339 Homo sapi |
| 756 | 16 | 2.7 | 25459 | 3 | SGAU6985 | Z67990 Caenorhabdi | 829 | 66591 | 67 | AC023031 | AC023031 Homo sapi |
| 757 | 16 | 2.7 | 25657 | 63 | AC014155 | AL006985 Streptomy | 830 | 67392 | 65 | AC018284 | AC018284 Drosophill |
| 758 | 16 | 2.7 | 27300 | 3 | SCD39 | AC014155 Drosophill | 831 | 67823 | 70 | AC026579 | AC026579 Homo sapi |
| 759 | 16 | 2.7 | 27887 | 3 | SCE50 | AL392146 Streptomy | 832 | 68334 | 61 | AC010563 | AC010563 Drosophill |
| 760 | 16 | 2.7 | 28519 | 93 | HSJ563E14 | AL117379 Human DNA | 833 | 68418 | 65 | AC018090 | AC018090 Drosophill |
| 761 | 16 | 2.7 | 29552 | 6 | CELF45F2 | U64845 Caenorhabdi | 834 | 68626 | 60 | AC006168 | AC006168 Drosophill |
| 762 | 16 | 2.7 | 31392 | 6 | LMFL5852 | AL499614 Leishmani | 835 | 68986 | 78 | AC091053 | AC091053 Homo sapi |
| 763 | 16 | 2.7 | 31720 | 93 | HSJ196E3 | Z68277 Human DNA s | 836 | 69986 | 78 | AC091053 | AC091053 Homo sapi |
| 764 | 16 | 2.7 | 32424 | 6 | CEZK131 | Z83245 Caenorhabdi | 837 | 69274 | 76 | AC084228 | AC084228 Homo sapi |
| 765 | 16 | 2.7 | 32610 | 65 | AC019840 | AC019840 Drosophill | 838 | 71512 | 94 | AF234171 | AF234171 Mus muscu |
| 766 | 16 | 2.7 | 33016 | 14 | SCCHXIVL | X86722 S. cerevisia | 839 | 71707 | 65 | AC018210 | AC018210 Drosophill |
| 767 | 16 | 2.7 | 33500 | 61 | AC011016 | AC011016 Leishmani | 840 | 72593 | 66 | AC021519 | AC021519 Homo sapi |
| 768 | 16 | 2.7 | 33818 | 3 | MTCY78 | Z77165 Mycobacteri | 841 | 73907 | 65 | AC018257 | AC018257 Drosophill |
| 769 | 16 | 2.7 | 34074 | 3 | SCF81 | AL13171 Streptomy | 842 | 75588 | 90 | AL356742 | AL356742 Human DNA |
| 770 | 16 | 2.7 | 34220 | 63 | AC014160 | AC014160 Streptomy | 843 | 76179 | 63 | AC012976 | AC012976 Drosophill |
| 771 | 16 | 2.7 | 34644 | 1 | AF080235 | AL359775 Leishmani | 844 | 76693 | 93 | HSJ475B7 | AL050306 Human DNA |
| 772 | 16 | 2.7 | 35042 | 6 | LMFL3665 | AL059837 Streptomy | 845 | 76748 | 94 | AC000400 | AC000400 Genomic s |
| 773 | 16 | 2.7 | 35437 | 3 | SCF43A | AL059837 Streptomy | 846 | 80167 | 65 | AC020038 | AC020038 Drosophill |
| 774 | 16 | 2.7 | 35722 | 63 | AC015427 | AC015427 Drosophill | 847 | 80327 | 72 | AC055889 | AC055889 Homo sapi |
| 775 | 16 | 2.7 | 35922 | 85 | AC000045 | AC000045 Homo sapi | 848 | 81662 | 12 | AB008265 | AB008265 Arabidops |
| 776 | 16 | 2.7 | 36148 | 60 | AC002321 | AC002321 Homo sapi | 849 | 83059 | 91 | AP001435 | AP001435 Homo sapi |
| 777 | 16 | 2.7 | 37130 | 3 | SCE9 | AL049841 Streptomy | 850 | 84056 | 91 | AC002121 | AC002121 Genomic s |
| 778 | 16 | 2.7 | 37799 | 6 | CELC05E11 | U53338 Caenorhabdi | 851 | 84702 | 12 | AB018114 | AB018114 Arabidops |
| 779 | 16 | 2.7 | 37840 | 85 | AC018301 | AC018301 Drosophill | 852 | 85095 | 4 | AC004573 | AC004573 Drosophill |
| 780 | 16 | 2.7 | 38000 | 67 | AC008974 | AC008974 Homo sapi | 853 | 85138 | 89 | AL137224 | AL137224 Human DNA |
| 781 | 16 | 2.7 | 38172 | 6 | LMFL5213 | AL583332 Leishmani | 854 | 85165 | 14 | NC99H12 | AL451018 Neurospor |
| 782 | 16 | 2.7 | 38702 | 4 | AC018661 | AC018661 Leishmani | 855 | 85402 | 92 | HS31B8 | AL033440 Human DNA |
| 783 | 16 | 2.7 | 38785 | 92 | HS695020B | AL049853 Human DNA | 856 | 85490 | 67 | AC022957 | AC022957 Homo sapi |
| 784 | 16 | 2.7 | 39329 | 85 | AC004091 | AC004091 Human Cos | 857 | 87210 | 15 | AB011485 | AB011485 Arabidops |
| 785 | 16 | 2.7 | 39741 | 3 | SC8E7 | AL591338 Streptomy | 858 | 87925 | 82 | AC005318 | AC005318 Homo sapi |
| 786 | 16 | 2.7 | 39796 | 9 | AX037570 | AX037570 Sequence | 859 | 88307 | 90 | AC009982 | AC009982 Drosophill |
| 787 | 16 | 2.7 | 39966 | 97 | HSU35G3 | Z23848 Human DNA s | 860 | 89301 | 92 | HS34B20 | AL031777 Human DNA |
| 788 | 16 | 2.7 | 40007 | 3 | SCAB10 | AL391515 Streptomy | 861 | 89794 | 12 | AC004665 | AC004665 Arabidops |
| 789 | 16 | 2.7 | 40118 | 97 | HSU63313 | U63313 Human cosmi | 862 | 91040 | 64 | AC017292 | AC017292 Drosophill |
| 790 | 16 | 2.7 | 40246 | 4 | AC024696 | AC024696 Caenorhab | 863 | 91480 | 93 | HSJ382F3 | AL079333 Human DNA |
| 791 | 16 | 2.7 | 40246 | 4 | AC024696 | AC024696 Caenorhab | 864 | 92854 | 93 | HSJ485C17 | AL049845 Human DNA |
| 792 | 16 | 2.7 | 40597 | 86 | AC007788 | AC007788 Homo sapi | 865 | 94268 | 61 | AC010046 | AC010046 Homo sapi |
| 793 | 16 | 2.7 | 41032 | 90 | AL354834 | AL354834 Human DNA | 866 | 94817 | 92 | HS981L23 | AL031686 Human DNA |
| 794 | 16 | 2.7 | 41807 | 3 | SC6D11 | AL158061 Streptomy | 867 | 96839 | 90 | AL357621 | AL357621 Human DNA |
| 795 | 16 | 2.7 | 41936 | 91 | CH19R30879 | AD000684 Homo sapi | 868 | 97698 | 65 | AC019670 | AC019670 Drosophill |
| 796 | 16 | 2.7 | 41971 | 4 | AC003765 | AC003765 Drosophill | 869 | 98572 | 68 | AC024573 | AC024573 Homo sapi |
| 797 | 16 | 2.7 | 42776 | 92 | HS11IA3 | Z79959 Human DNA s | 870 | 98572 | 68 | AC024573 | AC024573 Homo sapi |
| 798 | 16 | 2.7 | 43177 | 6 | CEIK06C4 | U64843 Caenorhabdi | 871 | 99277 | 61 | AC010568 | AC010568 Drosophill |
| 799 | 16 | 2.7 | 43177 | 6 | CEIK06C4 | U64843 Caenorhabdi | 872 | 99547 | 12 | AC068197 | AC068197 Genomic s |
| 800 | 16 | 2.7 | 43373 | 8 | AF083221 | AF083221 Fugu rubr | 873 | 100000 | 91 | AP000008 | AP000008 Homo sapi |
| 801 | 16 | 2.7 | 43644 | 85 | AC004092 | AC004092 Human Cos | 874 | 100000 | 91 | AP0000149 | AP0000149 Homo sapi |
| 802 | 16 | 2.7 | 43880 | 15 | SPAC17G6 | Z29162 S.pombe chr | 875 | 100373 | 65 | AC019995 | AC019995 Drosophill |
| 803 | 16 | 2.7 | 44699 | 5 | CBRG26F13 | AC084536 Caenorhab | 876 | 100771 | 65 | AC019995 | AC019995 Drosophill |
| 804 | 16 | 2.7 | 44978 | 89 | AF241731 | AF241731 Homo sapi | 877 | 103773 | 61 | AC009283 | AC009283 Homo sapi |
| 805 | 16 | 2.7 | 45029 | 64 | AC017280 | AC017280 Drosophill | 878 | 102225 | 63 | AC013623 | AC013623 Drosophill |
| 806 | 16 | 2.7 | 45142 | 5 | AF217651 | AF217651 Drosophill | 879 | 102621 | 92 | DMBR43E12 | DMBR43E12 Drosophill |
| 807 | 16 | 2.7 | 45328 | 85 | AC002128 | AC002128 Human DNA | 880 | 103268 | 92 | HS67M12 | AL138971 Drosophill |
| 808 | 16 | 2.7 | 45383 | 87 | AC010642 | AC010642 Homo sapi | 881 | 103568 | 92 | CNS07EGJ | AL590462 DNA cent |
| 809 | 16 | 2.7 | 47200 | 63 | AC013076 | AC013076 Drosophill | 882 | 104353 | 94 | HS117OK4 | AL022314 Human DNA |
| 810 | 16 | 2.7 | 49858 | 60 | AC008252 | AC008252 Drosophill | 883 | 104546 | 61 | AC010565 | AC010565 Drosophill |
| 811 | 16 | 2.7 | 49967 | 78 | AC090568 | AC090568 Homo sapi | 884 | 104725 | 66 | AC020928 | AC020928 Homo sapi |

| | | | | | | | | | | | | | |
|-----|----|-----|--------|----|--------------|---------------------|------|----|-----|--------|----|-----------|---------------------|
| 885 | 16 | 2.7 | 105160 | 61 | AC010571 | AC010571 Drosophila | 958 | 16 | 2.7 | 137599 | 72 | AC055855 | AC055855 Homo sapi |
| 886 | 16 | 2.7 | 105251 | 77 | AC084322 | AC084322 Mus muscu | 959 | 16 | 2.7 | 138615 | 60 | AC008139 | AC008139 Drosophila |
| 887 | 16 | 2.7 | 106661 | 60 | AC008855 | AC008855 Homo sapi | 960 | 16 | 2.7 | 138711 | 13 | AP003214 | AP003214 Oryza sat |
| 888 | 16 | 2.7 | 107159 | 92 | HS422R24 | AL031010 Human DNA | 961 | 16 | 2.7 | 139609 | 76 | AC080150 | AC080150 Homo sapi |
| 889 | 16 | 2.7 | 107732 | 75 | AC073794_3 | Continuation (4 of | 962 | 16 | 2.7 | 140031 | 61 | AC009246 | AC009246 Homo sapi |
| 890 | 16 | 2.7 | 108134 | 79 | AL157948 | AL157948 Homo sapi | 963 | 16 | 2.7 | 140920 | 88 | AC025173 | AC025173 Homo sapi |
| 891 | 16 | 2.7 | 108312 | 63 | AC013589 | AC013589 Homo sapi | 964 | 16 | 2.7 | 141207 | 93 | HS038210 | HS038210 Homo sapi |
| 892 | 16 | 2.7 | 108822 | 70 | AC027440 | AC027440 Homo sapi | 965 | 16 | 2.7 | 141991 | 84 | HS038210 | HS038210 Homo sapi |
| 893 | 16 | 2.7 | 108822 | 70 | AC027440 | AC027440 Homo sapi | 966 | 16 | 2.7 | 142114 | 12 | AC073166 | AC073166 Oryza sat |
| 894 | 16 | 2.7 | 110000 | 81 | AL391064_0 | Continuation (3 of | 967 | 16 | 2.7 | 142257 | 64 | AC017491 | AC017491 Drosophila |
| 895 | 16 | 2.7 | 110000 | 84 | HS171M_2 | Continuation (19 of | 968 | 16 | 2.7 | 142599 | 92 | HS413H6 | HS413H6 Homo sapi |
| 896 | 16 | 2.7 | 110000 | 84 | LMFLCHR12_18 | Continuation (4 of | 969 | 16 | 2.7 | 143159 | 87 | AL589986 | AL589986 Homo sapi |
| 897 | 16 | 2.7 | 110000 | 84 | LMFLCHR15_3 | Continuation (2 of | 970 | 16 | 2.7 | 143816 | 63 | AC013598 | AC013598 Homo sapi |
| 898 | 16 | 2.7 | 110000 | 84 | LMFLCHR32_01 | AL499622 Leishmani | 971 | 16 | 2.7 | 143888 | 62 | AC011687 | AC011687 Homo sapi |
| 899 | 16 | 2.7 | 110000 | 84 | LMFLCHR32_09 | Continuation (17 of | 972 | 16 | 2.7 | 143938 | 79 | AC073616 | AC073616 Homo sapi |
| 900 | 16 | 2.7 | 110000 | 84 | LMFLCHR32_26 | Continuation (20 of | 973 | 16 | 2.7 | 144092 | 79 | AL353146 | AL353146 Homo sapi |
| 901 | 16 | 2.7 | 110000 | 84 | LMFLCHR36_00 | AL499624 Leishmani | 974 | 16 | 2.7 | 144233 | 73 | AC068103 | AC068103 Homo sapi |
| 902 | 16 | 2.7 | 110403 | 72 | AC058813 | AC058813 Homo sapi | 975 | 16 | 2.7 | 144630 | 72 | AC064843 | AC064843 Homo sapi |
| 903 | 16 | 2.7 | 110403 | 72 | AC058813 | AC058813 Homo sapi | 976 | 16 | 2.7 | 144704 | 64 | AC016595 | AC016595 Homo sapi |
| 904 | 16 | 2.7 | 110990 | 83 | AP003408 | AP003408 Oryza sat | 977 | 16 | 2.7 | 144916 | 83 | AP003505 | AP003505 Oryza sat |
| 905 | 16 | 2.7 | 111865 | 79 | AL355883 | AL355883 Homo sapi | 978 | 16 | 2.7 | 145019 | 72 | AC053488 | AC053488 Homo sapi |
| 906 | 16 | 2.7 | 112460 | 92 | HS1104E15 | AL022312 Human DNA | 979 | 16 | 2.7 | 145331 | 86 | AC007877 | AC007877 Homo sapi |
| 907 | 16 | 2.7 | 112659 | 89 | AL137178 | AL137178 Human DNA | 980 | 16 | 2.7 | 145483 | 79 | AL162594 | AL162594 Homo sapi |
| 908 | 16 | 2.7 | 112958 | 61 | AC010107 | AC010107 Drosophila | 981 | 16 | 2.7 | 145744 | 14 | F414 | F414 Homo sapi |
| 909 | 16 | 2.7 | 113006 | 78 | AC090750 | AC090750 Mus muscu | 982 | 16 | 2.7 | 145803 | 77 | AC087110 | AC087110 Homo sapi |
| 910 | 16 | 2.7 | 113749 | 87 | AC008795 | AC008795 Homo sapi | 983 | 16 | 2.7 | 146275 | 83 | CNS01DPTX | CNS01DPTX Homo sapi |
| 911 | 16 | 2.7 | 114269 | 83 | AP002341 | AP002341 Homo sapi | 984 | 16 | 2.7 | 146275 | 76 | AC080119 | AC080119 Homo sapi |
| 912 | 16 | 2.7 | 114505 | 14 | F20P5 | 282210 Human DNA | 985 | 16 | 2.7 | 146923 | 77 | AC090197 | AC090197 Homo sapi |
| 913 | 16 | 2.7 | 118426 | 92 | HS431C21 | AC002025 Drosophila | 986 | 16 | 2.7 | 147021 | 86 | AC006501 | AC006501 Homo sapi |
| 914 | 16 | 2.7 | 118533 | 65 | AC020205 | AC008213 Drosophila | 987 | 16 | 2.7 | 147035 | 83 | AP003418 | AP003418 Oryza sat |
| 915 | 16 | 2.7 | 118603 | 60 | AC008213 | AL138742 Human DNA | 988 | 16 | 2.7 | 147259 | 66 | AC020845 | AC020845 Mus muscu |
| 916 | 16 | 2.7 | 118648 | 89 | AL138742 | AL138742 Human DNA | 989 | 16 | 2.7 | 147377 | 83 | AP003299 | AP003299 Oryza sat |
| 917 | 16 | 2.7 | 119853 | 89 | AL136374 | AL136374 Human DNA | 990 | 16 | 2.7 | 147416 | 86 | AC005549 | AC005549 Homo sapi |
| 918 | 16 | 2.7 | 121397 | 63 | AC012646 | AC002666 Drosophila | 991 | 16 | 2.7 | 147530 | 87 | AC008766 | AC008766 Homo sapi |
| 919 | 16 | 2.7 | 122114 | 60 | AC007573 | AC007573 Chlamydom | 992 | 16 | 2.7 | 147681 | 70 | AC027532 | AC027532 Homo sapi |
| 920 | 16 | 2.7 | 122395 | 78 | AC090433 | AC010496 Homo sapi | 993 | 16 | 2.7 | 148147 | 88 | AC044787 | AC044787 Homo sapi |
| 921 | 16 | 2.7 | 122672 | 91 | AP000704 | AP000704 Homo sapi | 994 | 16 | 2.7 | 148334 | 77 | AC087106 | AC087106 Homo sapi |
| 922 | 16 | 2.7 | 123288 | 91 | AP000547 | AP000547 Homo sapi | 995 | 16 | 2.7 | 148367 | 74 | AC069300 | AC069300 Oryza sat |
| 923 | 16 | 2.7 | 124182 | 78 | AC091087 | AC009187 Homo sapi | 996 | 16 | 2.7 | 148420 | 74 | AC069266 | AC069266 Homo sapi |
| 924 | 16 | 2.7 | 125041 | 73 | AC068653 | AC009372 Drosophila | 997 | 16 | 2.7 | 148455 | 92 | HS1114A1 | HS1114A1 Homo sapi |
| 925 | 16 | 2.7 | 125918 | 61 | AC009372 | AL137077 Human DNA | 998 | 16 | 2.7 | 148937 | 64 | AC015902 | AC015902 Homo sapi |
| 926 | 16 | 2.7 | 126712 | 89 | AL137077 | AC006574 Drosophila | 999 | 16 | 2.7 | 148937 | 64 | AC015902 | AC015902 Homo sapi |
| 927 | 16 | 2.7 | 127035 | 4 | AC006574 | AC087391 Homo sapi | 1000 | 16 | 2.7 | 148937 | 64 | AC015902 | AC015902 Homo sapi |
| 928 | 16 | 2.7 | 127211 | 77 | AC087391 | AC007983 Drosophila | | | | | | | |
| 929 | 16 | 2.7 | 128132 | 60 | AC007983 | AF241730 Homo sapi | | | | | | | |
| 930 | 16 | 2.7 | 128398 | 85 | AC004817 | AC020029 Drosophila | | | | | | | |
| 931 | 16 | 2.7 | 128428 | 89 | AF241730 | AC004241 Homo sapi | | | | | | | |
| 932 | 16 | 2.7 | 128469 | 65 | AC020029 | AC025473 Homo sapi | | | | | | | |
| 933 | 16 | 2.7 | 128473 | 85 | AC004241 | AC010453 Homo sapi | | | | | | | |
| 934 | 16 | 2.7 | 128540 | 69 | AC025473 | AC020761 Homo sapi | | | | | | | |
| 935 | 16 | 2.7 | 128937 | 61 | AC010453 | AL163539 Homo sapi | | | | | | | |
| 936 | 16 | 2.7 | 129290 | 66 | AC020761 | AC022414 Homo sapi | | | | | | | |
| 937 | 16 | 2.7 | 129800 | 79 | AL163539 | AC009170 Homo sapi | | | | | | | |
| 938 | 16 | 2.7 | 130149 | 88 | AC022414 | AP000679 Homo sapi | | | | | | | |
| 939 | 16 | 2.7 | 130948 | 61 | AC009170 | AP000679 Homo sapi | | | | | | | |
| 940 | 16 | 2.7 | 132639 | 91 | AP000679 | AC010059 Drosophila | | | | | | | |
| 941 | 16 | 2.7 | 132835 | 61 | AC010059 | AC079046 Homo sapi | | | | | | | |
| 942 | 16 | 2.7 | 132835 | 61 | AC010059 | AC017815 Drosophila | | | | | | | |
| 943 | 16 | 2.7 | 132930 | 75 | AC079046 | AC003999 Human PAC | | | | | | | |
| 944 | 16 | 2.7 | 133090 | 65 | AC017815 | AC064814 Homo sapi | | | | | | | |
| 945 | 16 | 2.7 | 133457 | 85 | AC003999 | AL442105 Oryza sat | | | | | | | |
| 946 | 16 | 2.7 | 133493 | 88 | AC064814 | AC004905 Homo sapi | | | | | | | |
| 947 | 16 | 2.7 | 134092 | 84 | HO702605 | AC011350 Homo sapi | | | | | | | |
| 948 | 16 | 2.7 | 134350 | 85 | AC004905 | AC037426 Oryza sat | | | | | | | |
| 949 | 16 | 2.7 | 134782 | 62 | AC011350 | AC037426 Drosophila | | | | | | | |
| 950 | 16 | 2.7 | 135056 | 12 | AC037426 | AP003078 Homo sapi | | | | | | | |
| 951 | 16 | 2.7 | 136056 | 60 | AC007571 | AC068279 Homo sapi | | | | | | | |
| 952 | 16 | 2.7 | 136493 | 83 | AP003078 | AC008278 Homo sapi | | | | | | | |
| 953 | 16 | 2.7 | 136555 | 73 | AC068279 | AC018412 Homo sapi | | | | | | | |
| 954 | 16 | 2.7 | 136868 | 86 | AC008278 | AC011243 Homo sapi | | | | | | | |
| 955 | 16 | 2.7 | 137267 | 65 | AC018412 | AP000598 Homo sapi | | | | | | | |
| 956 | 16 | 2.7 | 137269 | 87 | AC011243 | | | | | | | | |
| 957 | 16 | 2.7 | 137502 | 82 | AP000598 | | | | | | | | |

ALIGNMENTS

RESULT 1

LOCUS AR047920 3402 bp DNA PAT 29-SEP-1999
 DEFINITION Sequence 1 from patent US 5820864.
 ACCESSION AR047920
 VERSION AR047920.1 GI:5970263

KEYWORDS

SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 3402)
 Paranhos-Baccala,G., Lesenechal,M. and Jolivet,M.
 Trypanosoma cruzi antigen, gene encoding therefor and methods of
 detecting and treating chagas disease
 Patent: US 5820864-A I 13-OCT-1998;

FEATURES
 Location/Qualifiers
 source 1..3402

BASE COUNT 888 a 821 c 956 g 737 t
 ORIGIN

Query Match 100.0%; Score 594; DB 9; Length 3402;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 594; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|------|---|------|
| OY | 1 | caggtaacagcgttaacagcctttgtcttcaatctcgtacagccgaacggtatgctggctccgtgct | 60 |
| Db | 1232 | CAGGTACAGCCTTAACGGCTTTGCTTCAATCTGTAACGCGACGATAGCTGCTCTGGCT | 1291 |
| OY | 61 | gacatgtcgaattcgaattgaacgattccatctccgtgtccctccgcagggaaacacagcag | 120 |
| Db | 1292 | GACATGTGCAATTGCATTGAGGATCTTCATCTCCGGTCTCTCCGGCAGGGAACAACAGCAG | 1351 |
| OY | 121 | ccagggccaaaacatcgtgtagtgtgcgaacggaacccgggtgtgtgtcctcgggcaat | 180 |
| Db | 1352 | CCAGGCCAAAAAATCGTGTATGTGCGGACGCGCAAAACGGGGTGTTGTCTCGGGCACT | 1411 |
| OY | 181 | gaacggcgagatgacgaattacacaatgactctgcgcgtcgtgcaatccctctacata | 240 |
| Db | 1412 | GACCGCGCGATGTACACTCTATACCAATAGACATTCTGCGCTGTGTCATCCCTCATCA | 1471 |
| OY | 241 | ccccctgttcacgcgccagcgaagcagcgcgcgcctccctgcgcgacgagatcgcgttag | 300 |
| Db | 1472 | CCCCCTTTTACGGCCAGCCAAAGGACGCGGCGCTCTGCGCGCGCGATCGGGCTGAG | 1531 |
| OY | 301 | ccgcagcttgaggagcaagatcattgctaattctagtgaatcaagctggygattaatgltcaac | 360 |
| Db | 1532 | CCGCACCTGGGGAGCAAGATCATTTGCTAATCTGTGAATACACTGGGGAATTAAATGCACC | 1591 |
| OY | 361 | caaaggagcgtctgtacgacatgtgagcgcgcgcgcacagcagagttcaagcgtgtgacgtcc | 420 |
| Db | 1592 | CAAAAGGAGCGTCTGTACACTGTGAAGCGCGGCCACACAGAGCTTACGGCGGTGACGTCC | 1651 |
| OY | 421 | acgaactacgcgccgcagcgcgaacaatgcatacagggacaaatgggcgaactctgtacgct | 480 |
| Db | 1652 | ACGACTACCGCCCCCGCAGCGCAAAAGTCCATAGGGGACAAATGGCCGACCTTGACGGCT | 1711 |
| OY | 481 | ggattgtgtgcagctaaatagtgtgtgcacgcgcgcctgtcttcccaagcgcgcgcgaa | 540 |
| Db | 1712 | GGATTGTGTGCAGCTAATAGTGTGTGCACAGCGCGGCTCTGTCTCCACACCGCGCGAAA | 1771 |
| OY | 541 | ccaaacagagaagaagaagccttcgcgcgcgtgtgtgaacagactcctgttgcgata | 594 |
| Db | 1772 | CCAAACGAGAAAGAAAAGCGCTCTCGCGGCAATGTGAACACAGCTCCCTGTGGCGATA | 1835 |

| RESULT | 2 | | | |
|------------|---|------------|-----|--|
| LOCUS | A48910 | 3402 bp | DNA | |
| DEFINITION | Sequence 1 from Patent WO9605312. | | | |
| ACCESSION | A48910 | | | |
| VERSION | A48910.1 | GI:2302570 | | |
| KEYWORDS | | | | |
| SOURCE | Trypanosoma cruzi. | | | |
| ORGANISM | Trypanosoma cruzi | | | |
| REFERENCE | Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma; Schizotrypanum. | | | |
| AUTHORS | 1 (bases 1 to 3402) | | | |
| TITLE | Paranhos-Baccala, G., Lesenechal, M. and Jolivet, M. | | | |
| JOURNAL | NOVEL TRYPANOSOMA CRUZI ANTIGEN, AND GENE CODING THEREFOR; TH | | | |
| COMMENT | APPLICATION TO THE DETECTION OF CHAGAS' DISEASE | | | |
| FEATURES | Patent: WO 9605312-A 1 22-FEB-1996; | | | |
| source | BIO MERIEUX (PR) | | | |
| | Other publication CA 2173957 960222 | | | |
| | Other publication AU 3169195 960307 | | | |
| | Other publication FR 2723589 960216. | | | |
| | Location/Qualifiers | | | |
| | 1..3402 | | | |
| | ./organism="Trypanosoma cruzi" | | | |
| | /strain="G" | | | |
| | /db_xref="taxon:5693" | | | |
| | /dev_stage="PRIMASTIGONE" | | | |
| BASE COUNT | 889 a 818 c 958 g 737 t | | | |
| ORIGIN | | | | |

[illegible]

| RESULT | 3 |
|------------|--|
| LOCUS | TCU24190 |
| DEFINITION | TCU24190 3402 bp mRNA INV 04-AUG-1997 |
| ACCESSION | U24190 |
| VERSION | U24190.1 GI:790645 |
| KEYWORDS | |
| SOURCE | Trypanosoma cruzi. |
| ORGANISM | Trypanosoma cruzi |
| REFERENCE | Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma; Schizotrypanum. |
| AUTHORS | 1 (bases 1 to 3402) |
| TITLE | Lesenechal,M., Duret,L., Cano,M.I., Mortara,R.A., Jolivet,M., Camargo,M.E., da Silveira,J.F. and Paranhos-Baccala,G. Cloning and characterization of a gene encoding a novel immunodominant antigen of Trypanosoma cruzi |
| JOURNAL | Mol. Biochem. Parasitol. 87 (2), 193-204 (1997) |
| MEDLINE | 97391123 |
| REFERENCE | 2 (bases 1 to 3402) |
| AUTHORS | Lesenechal,M., Franco Da Silveira,J., Mortara,R.A., Duret,L., Camargo,M.E., Jolivet,M. and Paranhos-Baccala,G. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (06-APR-1995) Mylene Lesenechal, Biomerieux, 22 rue Saint Jean de Dieu, Lyon 69007, France |
| FEATURES | Location/Qualifiers |
| source | 1..3402 |
| | /organism="Trypanosoma cruzi" |
| | /strain="G" |
| | /db_xref="taxon:5693" |

```

misc_RNA      /dev_stage="eplmasticote"
               1..35
               /gene="Tc40"
               /note="spliced leader"
gene           1..3402
               /gene="Tc40"
               /gene="Tc40"
               /gene="Tc40"
               /codon_start=1
               /product="Tc40 antigen"
               /protein_id="AAC47657.1"
               /db_xref="GI:790646"
CDS            /translation="MTVTVDLENHAKPSNNEGRWMSVDAATENEVEAQRVLADSOFY
               /LAVTMKRRHVLRYVRKSNLKGTVRAHSPPIHAKVKVNRVSNVAASAGGEFVWVVT
               DETEASNGRPDLARLVYFKLQDPYIIPCSFPIINAEOSPDLLVYETQALID
               SSILIRFDVESLEATLORNCITLPTLTOPVSNSLCSYSGSGMFEETPEPVNACT
               LNRSTPSMAACEGEPVKALHLLDATTVEENSVLVAASKVQMLLTIVAPENLARK
               FYIDGSIYVMESSRETFAYFDDRKQALANMHSPIHFTCTHYMPCOVORNGCFNRT
               ADGSCVLADMSNRLLIFLRSRREDOQPOQKTSVATKPGCVSSGTAAASSSHNT
               TSAASAASPAAPVSPAPAKAAAPPAARSAEPHYGSKITIANLVNQLINTORSVSTG
               APATRTSTAVTSTTAPORTSPYGHNGRPVTAAGLVANSGASASPTAAKPTGEEK
               ASAACTSSVAIINATRALHNASLPQAPTDGYLAAVYQSEGEHOSLEBESVITNT
               SVYLKLLPPTIRDHOLNLIGLEAOMTELQOSRPORPORPTSSAKSSRPEVTYLY
               LIADLSRNITKCVKRVNDAITMHDHEVRHAIKRLQOTQNTIKSLDEKAKST
               TQPTAOLTOVENVLRLEAEVIGSLNGSLTSLVKNASLQELNSIMSSGVDEKRR
               MRELCITRESVAKRKRAITMDDSLHATSSFOGRRSAPEYTLATLSMVEEODROLE
               YMLMAOQPSILRLFLSLITRENENAYSELIENEFNDVWC SVLQDLIEAATEAKE
               VVGVVAIDILSERDQIAONGALGSKLITAMRAFERQARSETTSRSFLQCLIKLEKILLO
               S"
BASE COUNT    889 a      818 c      958 g      737 t
ORIGIN

```

```

Query Match      82.8%; Score 492; DB 96; Length 3402;
Best Local Similarity 99.7%; Pred. No. 4e-251;
Matches 592; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 caggtacagcgttaaggcttctgcttaacatcgtaacagcgaagtagctgcgtctcgtc 60
    |||||||
DB 1232 CAGGTACAGCGTAAACGCGCTTTGCTTCAATCGTACACGCCAGCGTACCTCGCTCGCT 1291

QY 61 gacatgctcgaatcgaatgaagatcttcacatctccggtcccccgaaggaagacagcag 120
    |||||||
DB 1292 GACATGTCGAATGATTCGACGATCTTCATCTCCGCTCTCCGACGAGAAACACACAG 1351

QY 121 ccagggcaaaaaaacatcgtgtagtggcgagcgaacccggggtgctgctcctcggcact 180
    |||||||
DB 1352 CCAAGGCCAAAAACATCGTAGTGGCGACGCGGAAACCGGGGTGTGTCTCTCGGCGACT 1411

QY 181 gaecggcgagtagcagtcataccaatacgaactctgcgctcgtcgtcattcccttgatca 240
    |||||||
DB 1412 GACGCGCGGAGTAGCACTCATACCAATACACATTTCTCCCTCGTCATCCCTGCAATCA 1471

QY 241 ccccccctttccagcgcaagcgacggcgctcctcctccggcgcgagcgaatcgagcggag 300
    |||||||
DB 1472 CCCCCTTTTTCACGCGCAGGAAAGCGCCGCGCTCTCTCCGCGGCGCATCGCGCTGAG 1531

QY 301 ccgcacgctgggggcaagatcattgtaatactagtagaactcagcttggggaattatgtcacc 360
    |||||||
DB 1532 CCCGACGTGGGAGCAAGATCATGTGAATCTAGTGAATCAGCTGGGGAATTAATGTACC 1591

QY 361 caaagaagcgtcgtcagcacttgagcgcgccgaagcgaaggtctacgcggtgaagctcc 420
    |||||||
DB 1592 CAAAGGAGCCTGTCACACTGGAGCCCGCCACGAGGAGGTCTTACGCGGCTACCTGCC 1651

QY 421 agcaactacgcccccgagcgaacaatgcatacgggacacatggcgaactctgtaagcgt 480
    |||||||
DB 1652 ACAGACTACCCCGCGGCGCAACAGTCCATACGGGACATAGCCACCTGTATACGGCT 1711

QY 481 ggaattgttgcaagctaatagttggtccagcgccctcgtctcccaacagcgcgcgcaaa 540
    |||||||
DB 1712 GGATGTGTGCGACTAATATGTGTGCGGCGGCGCTCGTCTCCACAGCGCGCGCAAA 1771

```

```

QY 541 ccaacaggaagaagaagcctccgcgcatgtgaaacgagctcgttgcgcata 594
    |||||||
DB 1772 CCAACAGGAGAGAGAAAAGCGCTCCGCGCATGTGTAACGACGCTCCGTGCGGATA 1825

RESULT 4
A48918 LOCUS A48918 22 bp DNA PAT 07-MAR-1997
DEFINITION Sequence 9 from Patent WO9605312.
ACCESSION A48918
VERSION A48918.1 GI:2202577
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 22)
AUTHORS Paranhos-Baccala,G., Lesenechal,M. and Jolivet,M.
TITLE NOVEL TRYPAOSOMA CRUZI ANTIGEN, AND GENE CODING THEREFOR, THEIR APPLICATION TO THE DETECTION OF CHAGAS' DISEASE
JOURNAL Patent: WO 9605312-A 9 22-FEB-1996;
COMMENT BIO MERIEUX (FR)
OTHER PUBLICATION CA 2173957 960222
OTHER PUBLICATION AU 3169195 960307
OTHER PUBLICATION FR 2723589 960216.
FEATURES
source location/Qualifiers
1..22
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 3 a      8 c      7 g      4 t
ORIGIN

```

```

Query Match      3.7%; Score 22; DB 9; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 cagcgaagtagctgcgtcct 56
    |||||||
DB 1 CAGCCGACGCTAGCTCGCTCCT 22

RESULT 5
AR047927 LOCUS AR047927 22 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 9 from patent US 5820864.
ACCESSION AR047927
VERSION AR047927.1 GI:5970270
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 22)
AUTHORS Paranhos-Baccala,G., Lesenechal,M. and Jolivet,M.
TITLE Trypanosoma cruzi antigen, gene encoding therefor and methods of detecting and treating chagas disease
JOURNAL Patent: US 5820864-A 9 13-OCT-1998;
FEATURES location/Qualifiers
1..22
/organism="unknown"
BASE COUNT 3 a      8 c      7 g      4 t
ORIGIN

```



```

NMA42491. NMA42491 331801 bp DNA BCF 04-DEC-2000
LOCUS      Neisseria meningitidis serogroup A strain Z2491 complete genome:
DEFINITION segment 4/7
ACCESSION  AL162755 AL157959
VERSION    AL162755.2 GI:7379742
KEYWORDS
SOURCE     Neisseria meningitidis Z2491.
ORGANISM   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria meningitidis Z2491.
REFERENCE  1 (bases 1 to 331801)
AUTHORS    Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,
            Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
            Davies,R.M., Davis,P., Devlin,K., Felwell,T., Hamlin,N.,
            Holtrop,S., Jagels,K., Leather,S., Moule,S., Mungall,K.,
            Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M.,
            Skellon,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.
            Complete DNA sequence of a serogroup A strain of Neisseria
            meningitidis Z2491
            Nature 404 (6777), 502-506 (2000)
JOURNAL    20222556
MEDLINE    2 (bases 1 to 331801)
REFERENCE  Parkhill,J.
AUTHORS    Direct Submission
TITLE      Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
JOURNAL    sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT    Notes:
            Details of N. meningitidis sequencing at the Sanger Centre are
            available on the World Wide Web.
            (URL, http://www.sanger.ac.uk/Projects/N\_meningitidis/).
FEATURES
SOURCE     location/Qualifiers
            1..331801
               /organism="Neisseria meningitidis Z2491"
               /strain="Z2491"
               /db_xref="taxon:122587"
               /note="serogroup: A"
            117..1333
               /gene="NMA1040"
               /gene="NMA1040"
               /note="NMA1040", possible type I restriction-modification
               system specificity protein, pseudogene, len: 1217 bp;
               N-terminus shows weak similarity to the C-terminal half of
               TR-050359 (EMBL:U25415), hsdS1B, Mycoplasma pulmonis
               restriction-modification enzyme subunit S1B (336 aa),
               fasta scores: E(): 2.3e-06, 26.2% identity in 141 aa
               overlap and to many hypothetical restriction-modification
               subunits. Also similar to NMA1041, fasta scores: E():
               4.8e-10, 37.0% identity in 100 aa overlap, C-terminus
               similar to part of SW:T1S1, ECOLI (EMBL:X13145), hsdS,
               Escherichia coli type I restriction enzyme Ecor1241I
               specificity protein (410 aa), fasta scores: E(): 1.9e-14,
               39.6% identity in 182 aa overlap. Similar to NMA1040, E():
               2.9e-09, 38.0% identity in 100 aa overlap. Contains Pfam
               match to entry PF01420 Methylase_S. Type I restriction
               modification DNA specificity domain. Contains a G(8) tract
               which would allow translation as an intact CDS, if
               variable. Lies within a region of unusually low GC
               content"
            /codon_start=1
            /pseudo
            /transl_table=11
            /product="pseudogene (putative type I
            restriction-modification system specificity protein)"
            complement(157..166)
            /note="Core DNA uptake sequence: gccgtctgaa"
            /label=DUS
            171..620
               /gene="NMA1040"
               /note="Pfam match to entry PF01420 Methylase_S, Type I
               restriction modification DNA specificity domain, score
               105.50, E-value 1.1e-27"
               /note="g(8)"
               726..735
                  misc_feature
                  /gene="NMA1040"
                  /note="Core DNA uptake sequence: gccgtctgaa"
                  /label=DUS
                  740..1231
                     misc_feature
                     /gene="NMA1040"
                     /note="Pfam match to entry PF01420 Methylase_S, Type I
                     restriction modification DNA specificity domain, score
                     158.70, E-value 1e-43"
                     1405..1408
                        RBS
                        1415..1409
                           gene
                           /gene="NMA1042"
                           1415..4509
                              CDS
                              /gene="NMA1042"
                              /note="NMA1042", pseudogene, probable type I
                              restriction-modification system restriction protein, len:
                              291 aa; similar to parts of many e.g. SW:T1B1, ECOLI
                              (EMBL:X13145), hsdR, Escherichia coli type I restriction
                              enzyme Ecor1241I R protein (EC 3.1.21.3) (1033 aa), fasta
                              scores: E(): 0.73.9% identity in 291 aa overlap (frame
                              1), followed by a stop codon, E(): 0.92.9% identity in
                              127 aa overlap (frame 1), followed by a frameshift near a
                              poly-A tract, E(): 0.72.5% identity in 604 aa overlap
                              (frame 3)"
                              /codon_start=1
                              /pseudo
                              /transl_table=11
                              /product="type I restriction-modification system
                              restriction protein (pseudogene)"
                              complement(1642..1651)
                              /note="Core DNA uptake sequence: gccgtctgaa"
                              /label=DUS
                              1687..1696
                                 misc_feature
                                 /gene="NMA1042"
                                 /note="Core DNA uptake sequence: gccgtctgaa"
                                 /label=DUS
                                 1761..1770
                                    misc_feature
                                    /gene="NMA1042"
                                    /note="Core DNA uptake sequence: gccgtctgaa"
                                    /label=DUS
                                    3906..3915
                                       misc_feature
                                       /gene="NMA1042"
                                       /note="Core DNA uptake sequence: gccgtctgaa"
                                       /label=DUS
                                       4514..4545
                                          stem_loop
                                          /note="Stem loop containing DNA uptake sequences: aaat
                                          gccgtctgaa at ttcagcgcg atttt"
                                          4519..4528
                                             misc_feature
                                             /note="Core DNA uptake sequence: gccgtctgaa"
                                             /label=DUS
                                             complement(4531..4540)
                                             /note="Core DNA uptake sequence: gccgtctgaa"
                                             /label=DUS
                                             complement(4559..6838)
                                                gene
                                                /gene="clpA"
                                                complement(4559..6838)
                                                   CDS
                                                   /gene="NMA1045, clpA, probable ATP-dependent protease
                                                   ATP-binding protein, len: 759 aa; similar to many e.g.
                                                   SW:CLPA, ECOLI (EMBL:M31045), clpA, Escherichia coli
                                                   ATP-dependent Clp protease ATP-binding subunit (738 aa),
                                                   fasta scores: E(): 0.56.2% identity in 762 aa overlap.
                                                   Similar to NMA1683, fasta scores: E(): 0.37.2% identity
                                                   in 844 aa overlap. Contains Pfam match to entry PF00495
                                                   clpA_B, Chaperonin clpA/B, PS00870 Chaperonins clpA/B
                                                   signature 1, PS00871 Chaperonins clpA/B signature 2 and
                                                   two PS00017 ATP/GTP-binding site motif A (P-loop)"
                                                   /codon_start=1
                                                   /transl_table=11
                                                   /product="putative ATP-dependent protease ATP-binding

```

protein" /protein_id="CAB84311.1" /db_xref="GI:7379743" /db_xref="SPTREMBL:O9JY15" /translation="MLSPEDHTLLIYLRARARAREPSTLEHLLVLIIEEAAVNV LKICGADLVKLSQALASVAVENTPQIPDHLLDTVEPTPIGFRVIRORAVHTOSGK AAEVPELDILVAHSESESHAVYFLKQSLITREYVLCIAGSDPDDGNDGLGEG EAEKQKTSLSDTVNLNNAEAGRIDPLGRHEMERLQILCRKKNPLVYGEAG VGTALAEGLAHIVKIGDIPDALKDAEVALDKGSLAGKRYGDEARVSVLKOLE KIPHALFEIDEHTITIGASGSGTGDASMLKLPALAKGALRCIGATTDEYRTFDK DHALSRFOKIDVEPTVEPTQILRGKMGEPGPHOVRYTQCALFAAEISARYNE RELPKDAIDVMDAGAORLPKRSKOKYIGKQIETVIKARIREKTYSHDQKYL OFGLRDKNNVIGQENALITLVAHMSRSGTLGPKPISFLSGPTGNGSKVQK LAYSMGVLPQREPMSEYMERHAYSLIGAPGVYGEQGLLAEVANKQPHCVLLDE IEKHAHPDIFVLLQVADAGKLTDNKGSADFRNVLIMTINAGASELSRPSLGTAKR ERGDEMOAIKINKLTPEPRNRDLAIIIPAPLSEPIYIVYDKFLQLEHRLDKVEAE FFPALRIAEKGEFDPMGARPNRRLIQEIRKIRPLADELLFGKLVGGCFRIMDMAK EEAIVKFKSKVPEPEYV"

misc_feature /gene="clpA" /note="Pfam match to entry PF00495 clpA_B, Chaperonin clpA/B, score 583.70, E-value 1.1e-171" complement(5228..5284)

misc_feature /gene="clpA" /note="PS00871 Chaperonins clpA/B signature 2" complement(5330..5353)

misc_feature /gene="clpA" /note="PS00017 ATP/GTP-binding site motif A (P-loop)" complement(5894..5932)

misc_feature /gene="clpA" /note="PS00870 Chaperonins clpA/B signature 1" complement(6173..6196)

misc_feature /gene="clpA" /note="PS00017 ATP/GTP-binding site motif A (P-loop)" 6374..6383

misc_feature /note="Core DNA uptake sequence: gccgtcggaa" /label=DUS complement(6842..7153)

gene /gene="NMA1046" complement(6842..7153)

CDS /gene="NMA1046" /note="NMA1046, len: 103 aa; unknown, similar to bacterial hypothetical proteins e.g. SW:YLJA_ECOLI (EMBL:AE00190), Y11A, Escherichia coli hypothetical protein (103 aa), fasta scores: E(): 7.5e-14, 44.6% identity in 83 aa overlap" /codon_start=1 /transl_table=11 /product="hypothetical protein NMA1046" /protein_id="CAB84312.1" /db_xref="GI:7379744" /db_xref="SPTREMBL:O9JY14" /translation="MNHPTDHSDDTLSDINTOPPKRYGVFLNDYTMERFVEIL TEVFMIAOEQAVAVMLLVHHEGKGLGCTYRDIAQFKQHVMERAKTEGHPKLCIYEE V"

Query Match 3.7% Score 22; DB 3; Length 331801; Best Local Similarity 100.0%; Pred. No. 0.58; Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 419 ccacgactacgccccgacgcg 440
DB 147264 CCACGACTACCGCCCCGACGC 147285

RESULT 7
AC011183 72056 bp DNA HTG 01-OCT-1999
LOCUS Homo sapiens clone 12_B_11, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC011183
ACCESSION AC011183.1 GI:6006118
VERSION
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 72056)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens, clone 12_B_11
JOURNAL Unpublished
2 (bases 1 to 72056)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Beldin, J., Barna, N., Beckerly, R., Boguslavsky, L., Bouhassira, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArillano, K., Dewar, K., Domingo, M., Donelan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heath, A., Horton, L., Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczy, J., Liu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.
COMMENT Direct Submission
Submitted (01-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
ALL repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html.
* NOTE: This record contains 80 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1
898: contig of 898 bp in length
gap of unknown length
899 1785: contig of 887 bp in length
gap of unknown length
1786 2678: contig of 893 bp in length
gap of unknown length
2679 3566: contig of 888 bp in length
gap of unknown length
3567 4474: contig of 908 bp in length
gap of unknown length
4475 5394: contig of 920 bp in length
gap of unknown length
5395 6335: contig of 941 bp in length
gap of unknown length
6336 7293: contig of 958 bp in length
gap of unknown length
7294 8172: contig of 879 bp in length
gap of unknown length
8173 9058: contig of 886 bp in length
gap of unknown length
9059 9962: contig of 904 bp in length
gap of unknown length
9963 10708: contig of 746 bp in length
gap of unknown length
10709 11617: contig of 909 bp in length
gap of unknown length
11618 12521: contig of 904 bp in length
gap of unknown length
12522 13412: contig of 891 bp in length
gap of unknown length
13413 14301: contig of 889 bp in length
gap of unknown length
14302 15220: contig of 919 bp in length
gap of unknown length

* 15221 16096: contig of 876 bp in length
* gap of unknown length
* 16097 17038: contig of 942 bp in length
* gap of unknown length
* 17039 17964: contig of 926 bp in length
* gap of unknown length
* 17965 18911: contig of 947 bp in length
* gap of unknown length
* 18912 19842: contig of 931 bp in length
* gap of unknown length
* 19843 20736: contig of 894 bp in length
* gap of unknown length
* 20737 21598: contig of 862 bp in length
* gap of unknown length
* 21599 22508: contig of 910 bp in length
* gap of unknown length
* 22509 23395: contig of 887 bp in length
* gap of unknown length
* 23396 24301: contig of 906 bp in length
* gap of unknown length
* 24302 25201: contig of 900 bp in length
* gap of unknown length
* 25202 26120: contig of 919 bp in length
* gap of unknown length
* 26121 27005: contig of 885 bp in length
* gap of unknown length
* 27006 27865: contig of 860 bp in length
* gap of unknown length
* 27866 28756: contig of 891 bp in length
* gap of unknown length
* 28757 29667: contig of 911 bp in length
* gap of unknown length
* 29668 30545: contig of 878 bp in length
* gap of unknown length
* 30546 31489: contig of 944 bp in length
* gap of unknown length
* 31490 32445: contig of 956 bp in length
* gap of unknown length
* 32446 33362: contig of 917 bp in length
* gap of unknown length
* 33363 34288: contig of 926 bp in length
* gap of unknown length
* 34289 35195: contig of 907 bp in length
* gap of unknown length
* 35196 36090: contig of 895 bp in length
* gap of unknown length
* 36091 36976: contig of 886 bp in length
* gap of unknown length
* 36977 37702: contig of 726 bp in length
* gap of unknown length
* 37703 38398: contig of 696 bp in length
* gap of unknown length
* 38399 39314: contig of 916 bp in length
* gap of unknown length
* 39315 40226: contig of 912 bp in length
* gap of unknown length
* 40227 41133: contig of 907 bp in length
* gap of unknown length
* 41134 42015: contig of 882 bp in length
* gap of unknown length
* 42016 42879: contig of 864 bp in length
* gap of unknown length
* 42880 43787: contig of 908 bp in length
* gap of unknown length
* 43788 44732: contig of 945 bp in length
* gap of unknown length
* 44733 45645: contig of 913 bp in length
* gap of unknown length
* 45646 46606: contig of 961 bp in length
* gap of unknown length
* 46607 47527: contig of 921 bp in length
* gap of unknown length
* 47528 48457: contig of 930 bp in length

* 48458 49401: gap of unknown length
* contig of 944 bp in length
* 49402 50300: gap of unknown length
* contig of 899 bp in length
* 50301 51179: gap of unknown length
* contig of 879 bp in length
* 51180 52130: gap of unknown length
* contig of 951 bp in length
* 52131 53074: gap of unknown length
* contig of 944 bp in length
* 53075 54010: gap of unknown length
* contig of 936 bp in length
* 54011 54906: gap of unknown length
* contig of 896 bp in length
* 54907 55797: gap of unknown length
* contig of 891 bp in length
* 55798 56687: gap of unknown length
* contig of 890 bp in length
* 56688 57579: gap of unknown length
* contig of 892 bp in length
* 57580 58443: gap of unknown length
* contig of 864 bp in length
* 58444 59345: gap of unknown length
* contig of 902 bp in length
* 59346 60251: gap of unknown length
* contig of 906 bp in length
* 60252 61197: gap of unknown length
* contig of 946 bp in length
* 61198 62066: gap of unknown length
* contig of 869 bp in length
* 62067 62960: gap of unknown length
* contig of 894 bp in length
* 62961 63871: gap of unknown length
* contig of 911 bp in length
* 63872 64769: gap of unknown length
* contig of 898 bp in length
* 64770 65714: gap of unknown length
* contig of 945 bp in length
* 65715 66596: gap of unknown length
* contig of 882 bp in length
* 66597 67499: gap of unknown length
* contig of 903 bp in length
* 67500 68408: gap of unknown length
* contig of 909 bp in length
* 68409 69331: gap of unknown length
* contig of 923 bp in length
* 69332 70264: gap of unknown length
* contig of 933 bp in length

Query Match 3.4%: Score 20; DB 62; Length 72056;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 538 aaaccacaggaagaagaaa 557
|||||
Db 39898 AAMCCAAAGAGAGAGAGAAA 39917

RESULT 8
AC011438/c
LOCUS AC011438
DEFINITION Genomic sequence for Arabidopsis thaliana BAC T23G18 from
AC011438
ACCESSION AC011438.3 GI:6289106
VERSION AC011438
KEYWORDS HTG.
SOURCE
ORGANISM Arabidopsis thaliana
thale cress.
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 91470)
Shinn,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C.,

TITLE
JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 91470)
Ecker,J.R.
Direct Submission
Submitted (07-OCT-1999) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
(bases 1 to 91470)

REFERENCE
AUTHORS

Submitted (09-NOV-1999) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
(bases 1 to 91470)

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (15-DEC-1999) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th and
Hamilton Walk, Philadelphia, PA 19104-6018, USA
(bases 1 to 91470)

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th and
Hamilton Walk, Philadelphia, PA 19104-6018, USA
(bases 1 to 91470)

REFERENCE
AUTHORS

Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C.,
Shinn,P., Altati,H., Bel,B., Chin,C., Chlou,J., Choi,E., Conn,L.,
Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B.,
Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharsky,N.,
Nguyen,M., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J.,
Southwick,A., Thaverl,A., Toriumi,M., Vaysberg,M., Yu,G., Davis,R.,
Fedspsiel,N., Theologis,A. and Ecker,J.

Direct Submission
Submitted (11-OCT-2000) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th and
Hamilton Walk, Philadelphia, PA 19104-6018, USA
On Nov 9, 1999 this sequence version replaced gi:6139057.

COMMENT
FEATURES

Location/Qualifiers
I. 91470
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="T23G18"
complement(join(831..899,1015..1092,1195..1329,1360..1473,
1513..1586,1687..1758,1838..1899,2653..2775,2882..3013,
3472..3559,3638..3732,3827..4057,4288..4404,4481..4582,
4711..4827,4948..5733))
/note="putative DNA ligase spl042572"
/codon_start=1
/evidence=not_experimental
/product="T23G18.1"
/protein_id="AAFI8258.1"

/db.xref="GI:6579215"
 /translation="MLAIRSSNNYLRCPISLCTKTKIOGFSVLISFROSLIRLBSGCC
 HRAMSSSRPSAFDALMSNARAARKKTPQTTNLSRPNKRKIGSTQDANKGTTVSGC
 TLPEKEDLELIVSDSANPRSDTSSIAEDSGAKVAKTSLTSDMKSKTGLKRLKVD
 LPEKMSKWEKGERVPELVALAFDLISNEGRIAYTIDLCNMLRYIATTEKLPVND
 VYLSNEMETAPNHEGVLEIGESTIIKALISEAFGEDEHVKKONTLEDGLIVAGSSGS
 TOTMNEKPEPLTVVPEPTFOPIAKSESDNEKKNNKMLALVATTDCELYTLLRLL
 OAKLISGSGGVIALALGOAAVYVNEEHSKPPNNKPSLEBAKIVKOVFLPVLYDITI
 VPALLSGGVNMLPKTCQNTFLTVPICPGIMLAKTCQNTVAETLKKFOIVYVECEKYGEEA
 OIHHEDEDTFELISSMAERNKTFPVALASRLKRSYVSFLIDCEVYAFREKKITL
 LPQLLSTRARNVAVNDIKVGCITFAFDMLVNGQOLIOENLIRREKLTESDEED
 GYFOATVATNSNDIDELQFLDASVDVIGSDVYIAVAHNGKRTGYGVALGALV
 DVLNDEEEOQSLCKIGTGFSDAMLDERSSLSQVIAIPKHEFLTNCDSNIOQYRVAG
 DYNDKWEPEVEVGLCLFRHSIYIVVMKMKKTCVCLKCPKHEVFKADTLISVHIAATG
 IYDPKGGSLNPEFLRLRYREDKKPPEEAATSSQINDIYOAKOHNHNSVEVGGDD"
 /comment["(oln(6737, 7168, 7256, 7909, 8003, 9522,
 10196, 10618, 10842, 11504, 11599, 12614, 12710, 12962,
 14353, 14725, 14825, 15484, 15589, 15608, 15666, 17013,
 17655), 18255)"]
 /note="putative Na/H antiporter gl14063740; similar to
 ESTS gbl141634, gb1AA586073"
 /codon_start=1
 /evidence="not experimental"
 /product="123618.2"
 /protein_id="AAE18257.1"
 /db_xref="GI:6579214"
 /translation="MVPPNPAHOPARRTPOLQDQSPRAQPLPGRANPVLCITVAL
 VLGLGVALITLITVTLRPRILYTVAAASQVFGALGNDDHINAESYIKENNEE
 KHVSRYSRSLSTAHNDSVAHKHISFQKPRNNEETIQLSHVNALEAFNAROD
 BAERKSGITMEVTVTATVASYKKTWIPERRRTKATCPCKKNLCHMSNREAEVAY
 NATFGDSYHNGYCGSEDSKCPDQPLVANSFGWEKLDPIHGKNTIWDSPHLES
 YIVLALQKQFVYLSLKGIGLVPKRTISSMLVRLTSLSSFGAALISVTLNLPNDT
 ILFPDTPKAPVETLGLGFAFVFEVCEVKNVAKVNTKQVLTQIATVILDI IAA
 NMVPEKLEETGGKYLTMGEYETILFMQOISAFETGSLRDLRDLSEFGVITVIA
 VADGQFPEVETLVAAMDMSRYALOGAGIIGIYFEMVYVAPPMFVFKTPOEDRV
 KCEPIYIILIAFGGYFLKEIHMPNAPGLDGLVCHPGSLQDLEKREKSTPGLV
 LILPLFMSLQIOGPMYLANIGLRHEDQOYALRTIIVYPAKILFMSYALLAK
 MFLDSEFMALISKGIYELCYFLTYGVSNAVAKST IAAITLVISTISPLIHR
 LLDSSKRTFQKRLMSKLGSSELCTVCIHKDHLISGMLNLAQSPHLESTISCC
 YIHVELVGLNDPVIISHQOKAEBKNSYSNNVLAEDNKHWSISLELFTICISN
 PFMHOEITSLALDQASFEALPRHIIMSLDQTVSDVWRANRANVLYLAQCSYGS
 IYVHOKIETSLAOKSSPEVCAIFVGGKDDDEALAGQMMRANVNLVYKILPAKMK
 IYVHOKIETSLAOKSSPEVCAIFVGGKDDDEALAGQMMRANVNLVYKILPAKMK
 DGMTGMOMLDQAEVKEVLRNNNTVGOHSEFVDEETVNDGSDTSLIISINSEDF
 LEVYRSAGVGTDAVYSALSEWDEDELEIDMSYVDEWCKGEPENSSITGESPHH
 VNSHIMEVMTFRKGMNWEYPLMLLELTISTFPIHRLDISNKGILRVPRTYMM
 IVRLILGOTCHPSNKHIDHIFPDNDNPKAETLGAEPVLYMLGVTMDALPEV
 TERKSSVGTIVYIPLIGSLGTYFERRRODRLNRAEYLIIFLOSISAFETIDIL
 LKLDQIKSESERIALSGAMVDTMLAEVTFPMAIYKELTGEMQVGFCLFVMMKIL
 VYRPAMYIYKOTPGRPVNDYELISIGIAPFACPTFNKVIYHLEPGASVFCGLVAV
 NGYLGITLLQTFESFNLSGILPLGSLITMNOVDLRIKESGDLIMEGOIYVISEFS
 ILLVNTTFYVTTITAYAFKAPLRBSPALAYLSKGFELAEYTYVAELIRPEVET
 TLLAVTTLNSTFPLMLELTVHDPKRCYRKYRNLGLKSGAALQCLMCEYRDHITIT
 SMTDLEFFSPQSDSPACNCLILHVELVGLNAMPITSHQLOPEGSSYSDNYIISF
 RGFOQFPEYSLDIFTSVSYSOHNHEDGQMLALSRSLSLIVPEFHRTWSDRSTVY
 NDDNRLMTANVNLVLRAPCSVGIYFYRKRIYVSHMAKSHKICLLIFNGKDKRELA
 NMRLTERTKRTLIIRFIPKSEMDNDEMOQSNLKESETSYVSGNKKDKAKVATY
 IDKAVSDSESRSLIRAMANDYDLFVSSGGIGRAEPLISWMTPEHNEGIDTLAIA
 SHEYRSKILMDCEITRPVLLVILFTAKSHCOPEPLPIYELIRODLSOMDAENATV
 REKEMNNDERKAEMGTWKFCDVSPHMLNSHGAERKASSKGDREYLPOLEIIL
 ILISLELRDPMLEKRLGVPITPKFTSMILVAVNQGIPISACWHMCSNNITHELEFL
 VGALISEWFGSKOIPCLKHIFHYNQWTKYPLDTIGAFALNDELRYVDTQIMKRL
 SGTCSVUIGTSMIPQOIGKLTLSRSKRSILITMEYETVMTFMSMPEFCVMMKIL
 LTDLCTVHRDFOILOSAGVNTDLAFLPILVAYSVARETQVGMGLAFMAFLEFVIL
 VROFLMYLIRHPRBGAQVKNYKVIYGLLALVLYMSRPLPFGICAFGLAVPNQD
 PPLGSVFLQKDESNEGFLPDLFGLSLNLIKDMSEFLKREFGNRLHGMHCEGSPIL
 IYVIAKFAISPLAALATKIPLDSDIILVINGTSSSEFLGVILVLAPEKDRISLEVISL
 LGVYILVNSLITLPPAIFRLYDRSKRFVYGRHNLKEREKEMOTLCKINKPDIITMSI
 LEATSPDSPEECVCLHILDELGOATPTFISHROKCPKPSRSISEVSISSFOEDB
 YVWDSASINMFLSTISAKHEMOICWFLSOGSNILISFRMEDEKANTVILDDOTUL
 RSLNNTLKRAKPSVGIYFYRKRIYVSHMAKSHKICLLIFNGKDKRELAENATV
 NOOVILTVLRLPSTYADDESLRIHSQVYDANKRHEDORPKGSTIIDTYVGDGTETSK
 ILHASYDYDLFIVGRBSGVGTVTYRGIGDMWNEPELAVIGDILLASEFPBSRASVAVV
 000E"

CDS

complement(20859..21590)
/note="putative histone H2B-III sp|P54346"

/codon_start=1

/evidence=not_experimental

/product="T23618.3"

/protein_id="AAFI8256.1"

/db_xref="GI:6579213"

/translation="MAPRRPKVSVTKKKKVEETIKYTVTEEGDPCVITETANDOE

ODLFESTVPGENNTVEIPVEPDESLPGENETTVKIPVDDDESSPOPEPVEV

RDESPPEPPEPRAKSGTLTKTDKVEKKDKKKKKKKKKDDIAGDEYRIVYKVK

QHPDLCITSKAMTVNMFGDMFERIAQEAALSDYTKRTKRTSSRELEAAVRLVPS

ELSRHAAEGSKAVSNFYDSSRR"

complement(22659..22994)

/note="hypothetical protein"

/codon_start=1

/evidence=not_experimental

/product="T23618.4"

/protein_id="AAFI8255.1"

/db_xref="GI:6579212"

/translation="MSKLETLDEEKTVEOKPRSOEEDHODSSKKEELLESCTPS

SDHKIPEVCCPPRRRRPREISITKTRLSKDLRFPEADVGSQEVETLVEHPENH

RKKRNSA"

join(25370..25804,26086..26363,26430..26609,26706..26963,

27066..27198,27287..27377,27474..27576,27937..28040,

28126..28219,28391..28564,28626..28688,28796..28861,

29183..29410,29502..29579,29668..29799,29916..30056,

30110..30195,30414..30495,30584..31060)

/note="putative vacuolar assembly protein VPS1 homolog

sp|P93331: similar to ESTs gb|AA042360, gb|I20891"

/codon_start=1

/evidence=not_experimental

/product="T23618.5"

/protein_id="AAFI8239.1"

/db_xref="GI:6579196"

/translation="MAVPPENGVGDGEREEDEDEEEDEEEDEEENGDEAEPEPL

KYOMGNVALLSDNASCIAVAARMAIALGTHDFRAHTAPVNDINDTGEYIGSC

SDDSVYINSLFTDDEKMKEDYHRPKAISLDPYTKQSKRFAGAGHLNYSK

WFGNKDQVCENVLTGEGEPHISVKGCSLIAMANDGVKVDYKDDGVTFIEKPS

PRPALPHLYWODDTLLVIGMGTSVKIASIKSDOOGTGFROIOMSLTNOVDIVAP

OTSYTISGIAPGDSLIYAVIPLEGDEKESSTTSLRGNMORPRTIRVSNND

LMDALPVHGEFHKAKDYSLAAHPFGSSVAGQMAAGDEPLYIYSPKDVYIAKPR

CDS

Query Match 3.4%; Score 20; DB 12; Length 91470;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 cttctgcgtcgtcgtcgtcc 231

Db 42747 CTTCTGCCGTCGTCGATCC 42728

RESULT 9

AC017643

LOCUS

DEFINITION

AC017643 177816 bp DNA

AC017643

VERSION

KEYWORDS

SOURCE

ORGANISM

AC017643

AC017643.1 GI:6554354

HTG; HTGS_PHRASE2.

fruit fly.

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 177816)

Adams,M. and Venter,J.C.

Submitted (10-DEC-1999)

Rockville, MD, USA

This sequence was identified as CDM-10211585 by the submitter.

For more information on this record e-mail to fly@celera.com.

* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced

FEATURES

Source

BASE COUNT

50527 a 38300 c 39085 g 49904 t

ORIGIN

1..177816

/db_xref="taxon:722"

/organism="Drosophila melanogaster"

Query Match 3.4%; Score 20; DB 65; Length 177816;

Best Local Similarity 100.0%; Pred. No. 7.6;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 gctctgcgtcgtcgtcgtc 69

Db 164707 GCGTCTCGTCGTCGATCG 164726

RESULT 10

AC007697

LOCUS

DEFINITION

AC007697

AC007697

VERSION

KEYWORDS

HTG.

fruit fly.

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 194897)

Celniker,S.E., Adams,M.D., Krommiller,B., Tyler,D., Wan,K.H.,

Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,

Rogers,Y., An,H., Baldwin,D., Bazon,J., Beeson,K.Y., Busam,D.A.,

Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,

Dodson,K., Dorsett,E., Doup,L.E., Doyle,C., Dresnek,D., Fartan,D.,

Ferreira,S., Frisoe,J., Galle,R.F., Garg,N.S., George,R.A.,

Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,

Ibegwam,C., Jalali,M., Kruse,D., Li,P., Matell,B., Moshrefi,A.,

McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,

Pacled,J., Paragay,V., Park,S., Patel,S., Pfeiffer,B.,

Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,

Stapleton,M., Strong,R., Svirskaas,R., Tector,C., Williams,S.M.,

Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.

Sequencing of Drosophila chromosome 2R, region 54D4-54E7

Unpublished

2 (bases 1 to 194897)

Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazek,R.G.,

Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,

Doyle,C.M., Fartan,D.E., Galle,R., George,R.A., Harris,N.L.,

Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,R., Kearney,L.,

Kim,E., Lee,B., Lewis,S., Li,P., Lomtan,M.A., Madada,P.,

Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacled,J.M., Park,S.,

Pfeiffer,B., Poon,L., Segueira,A., Sethi,H., Sibir,E.,

Svirskaas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and

Rubin,G.M.

Direct Submission

Submitted (02-JUN-1999)

Submitted (02-JUN-1999)

Submitted (02-JUN-1999)

Submitted (02-JUN-1999)

Submitted (02-JUN-1999)

Submitted (02-JUN-1999)

Submitted (02-JUN-1999)

Submitted (02-JUN-1999)

Submitted (02-JUN-1999)

Submitted (02-JUN-1999)

Submitted (02-JUN-1999)

| Query Match | Best Local Similarity | 3.4% | Score 20 | DB 4 | Length 194897 |
|-------------|---|--------------|----------|-------------|---------------|
| Matches 20 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 | |
| QY 50 | gcgtctgcgtgcacatgcgc 69 | | | | |
| Db 128453 | GCgcTCTGCGTGCACATGTCG 128472 | | | | |
| RESULT 11 | | | | | |
| AE003802 | 262395 bp | DNA | INV | 04-OCT-2000 | |
| LOCUS | Drosophila melanogaster genomic scaffold 14200013386047 section 41 | | | | |
| DEFINITION | of 52, complete sequence. | | | | |
| ACCESSION | AE003802 AE002787 | | | | |
| VERSION | AE003802.2 GI:10727480 | | | | |
| KEYWORDS | HTG. | | | | |
| SOURCE | Fruit fly. | | | | |
| ORGANISM | Drosophila melanogaster | | | | |
| REFERENCE | Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. | | | | |
| AUTHORS | 1 (bases 1 to 262395) Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amaratunga,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Yeung,M.D., Zhang,Q., Chen,L.X., Brandon,R.C., Rogers,Y.H., Blazey,R.G., Champe,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor Miklos,G.L., Abtill,J.F., Agbayani,A., An,H.J., Andrews-Plamkoc,H.C., Baldwin,D., Bailey,R.W., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y., Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borikova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burris,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferrara,C., Ferriera,S., Fleischmann,W., Foster,C., Gabrielian,A.E., Gary,N.S., Gelbart,W.M., Glasser,K., Gloeck,A., Gong,F., Gorrell,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D., Helman,T.J., Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wei,M.H., Ibegwam,C., Jalili,M., Kalush,F., Karpen,G.H., Ke,Z., Kienison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Matel,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Mekukolu,G., Milshina,N.V., Mobarry,C., Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nusskern,D.R., Paclob,J.M., Palazolo,M., Pittman,G.S., Pan,S., Pollard,D., Puri,V., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shie,B.C., Siden-Kiamos,I., Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C., Stapleton,M., Strong,R., Sun,E., Svirskis,R., Tector,D., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassenaar,D.A., Weinstock,G.M., Weissbach,J., Williams,S.M., Woodage,T., Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yen,R.F., Zaveri,J.S., Zhang,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., | | | | |

```

TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
The genome sequence of Drosophila melanogaster
Science 287 (5461), 2185-2193 (2000)
2 (bases 1 to 262395)
Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
Direct Submission
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
On Oct 9, 2000 this sequence version replaced gi:7302706.
Location/Qualifiers
1..262395
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="2R"
<1..27
/feature="CG4878"
/product="CT43034"
/db_xref="FLYBASE:FBan0004878"
/db_xref="FLYBASE:FBgn0034237"
<1..27
/feature="CG4878"
/product="CT15671"
/db_xref="FLYBASE:FBan0004878"
/db_xref="FLYBASE:FBgn0034237"
<1..27
/feature="CG18432"
/product="CT41966"
/db_xref="FLYBASE:FBan0018432"
/db_xref="FLYBASE:FBgn0034238"
/feature="not_experimental"
<246..>2757
/feature="CG18432"
/map="54C10-54C12"
/db_xref="FLYBASE:FBan0004878"
/db_xref="FLYBASE:FBgn0034237"
join(<246..1745,1804..>2757)
/feature="CG18432"
/product="CT41966"
/db_xref="FLYBASE:FBan0018432"
/db_xref="FLYBASE:FBgn0034238"
/feature="not_experimental"
<246..>2757
/feature="CG18432"
/map="54C10-54C12"
/db_xref="FLYBASE:FBan0018432"
/db_xref="FLYBASE:FBgn0034238"
/feature="not_experimental"
join(246..1745,1804..2757)
/feature="CG18432"
/note="CG18432 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0018432"
/db_xref="FLYBASE:FBgn0034238"
/feature="not_experimental"
/protein_id="AAE57841.1"
/db_xref="GI:7302763"
/translation="MEALASRKTKAKKKRKTHTTGRYKRRDEYOKLIRATEGRISGE
HDEMLASLQFFHGFGEQMOENDFAVEDCSDSDSGELAEPTDSQSQTFIHIDLH
DPLPLAMQLFKORHYDESLLYRPSNTLHFQLHLSDELKDSQIRMLNRYQSQSKS
SGRKSLGILGOLSGOKASPLFMGKHLYRSYCHRRFQSGVFPMPSPFLKLPBRCKL
LCIKOLRPSLHROLLEHRIAROLLEDFQLEYSOORORICRKLRELEETAROVAKILT
ASGQODTAEVKKLKTROLROKRYAESAAQRNLROLLKEMAKLKELRQORFQCT
RFQGLRVHHPDLLEASYCAMKSEFFEDLAEVYREHLELTYTRLRMTDOKSHRSAT
GHSPPKRPQDFRIMASLRKYEDKTDPEPYEVFRLHADATARNLSTPGDOLPK
ARNVFLTKFLDGOFGVQSRTRYRLEPFDQISINBCIGVLLERSLPEMLNINILYKSTLT
PKSRRIALISTPLVLSKSDAVOEKSLFQSLAGADVYLVEYVSTEGMSALHL
DDVORLPDALRQILPOSTLPALPOAKSEIVLTPRPSGRIRKKSQPLSLIFAEOLQSG
LDVILNRRFOLLHSRHOQNLNLTQKQRYPALEOEISSEEPPLPDGCTVQLLEPGTY
WNPFLDKLHGRKFLQILVEYIASQSRARAKALQTPLEPLLLLEDSDRSGTGMTALM
RARCSPYQGOGRNLSLRPSGSMGHSAPDLFKSFCSVLHVVRATGVPVRSRHLINLE
RRSAGGDMSSTSLFVTOSE"
join(<2926..3204,3268..3993,4054..4292,4416..>4506)
/feature="CG18631"
/product="CT41968"
/db_xref="FLYBASE:FBan0018631"
/db_xref="FLYBASE:FBgn0034239"
/feature="not_experimental"

```

```

gene
<2926..>.4506
/gene="CG18631"
/map="54C10-54C12"
/db_xref="FLYBASE:FBan0018631"
/db_xref="FLYBASE:FBgn0034239"
/evidence="not_experimental"
join(2926..3204,3268..3393,4054..4292,4416..4506)
/gene="CG18631"
/note="CG18631 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0018631"
/db_xref="FLYBASE:FBgn0034239"
/evidence="not_experimental"
/protein_id="AA57840.1"
/db_xref="GI:7302762"
/translation="MYSNVRPFYTLSTGRLCRSRTAECSNPTMNEOLOIQOSFGD
LRDLKLTSLDELIEQYSDASDVYQVOCNMIGERRVPIINSLLASRTREGCIELAM
PKVLVGRPLIDSVTMPTDQYPEFKEAVHLMFYLSEPGGDLAPLPCALCAEK
PELQOYDERLELQQLLPQRYVEPLCTAQCGRVCLTRIMDPVLPRAVSGEN
PLEICIRYSLSLHSRSDPCGFRGVMNLSLDTWCSKVDGLVLCNMYMSTGL
ECWLIIGVSCPHGCSFVLRQPELALFLVSPATGRKRYOLOPHCLRRICVYVGO
NIYFNITETRVSKTHFNLDGACMFLFRVRVAPDSGVOKLDYAKRSLESLQOK
HIERKIMRKISAMRTTKTWMNRGRYDEPLFSEQLEREFPYSLAIFRK"
join(6206..6413,6478..6861,6918..7090,7164..7308)
/gene="cyp33"
/note="Nucleotide sequence of the Celera sequence differs
from the published sequence for this transcript."
/product="CT1569"
/db_xref="FLYBASE:FBan004886"
/db_xref="FLYBASE:FBgn0028382"
<6206..>.7308
/gene="cyp33"
/note="CG4886"
/map="54C11-54D1"
/db_xref="FLYBASE:FBan004886"
/db_xref="FLYBASE:FBgn0028382"
join(6213..6413,6478..6861,6918..7090,7164..7308)
/gene="cyp33"
/note="cyp33 gene product; Nucleotide sequence of the
Celera sequence differs from the published sequence for
this transcript"
/codon_start=1
/db_xref="FLYBASE:FBan004886"
/db_xref="FLYBASE:FBgn0028382"
/protein_id="AA57839.1"
/db_xref="GI:7302761"
/translation="MSNDKRTIYVGLADEYTERLNNAPFPGDIADIQMPADYESQ
RHKGFATIEYQSEDAANAIDNMNDSLCRTIRVNLAKEPVKEDSFKPIWADDML
OKHAGATLQEGEPEAEKAVETPTGPAVIEKAERKNQVFEIDIRGNDAGRIMLAR
ADVVPKTAENRQICTHEOGYKGCSEFHVIPPEMCGGFTNNNGTGSGIKKKE
NDENFNLKHSFGLSMANGSANGMSQFICTTMTMDLNKHVYFGHVISGAIVYRK
MERGSKSGSPSKIVYISCELEK"
complement(join(7324..7910,8333..8405))
/gene="Rpl18A"
/note="Nucleotide sequence of the Celera sequence differs
from the published sequence for this transcript."
/product="CT20271"
/db_xref="FLYBASE:FBan0006510"
/db_xref="FLYBASE:FBgn0010409"
complement(<7324..>.8405)
/gene="Rpl18A"
/note="CG6510"
/map="54C11-54D1"
/db_xref="FLYBASE:FBan0006510"
/db_xref="FLYBASE:FBgn0010409"
complement(join(7395..7910,8333..8350))
/gene="Rpl18A"
/note="Rpl18A gene product; Nucleotide sequence of the
Celera sequence differs from the published sequence for
this transcript"
/codon_start=1
/db_xref="FLYBASE:FBan0006510"
/db_xref="FLYBASE:FBgn0010409"

```

```

/protein_id="AA57838.1"
/db_xref="GI:7302760"
/translation="MRAGLKEEYVGRKIPSEKQPTPIYKRIRAPDNIYAKSRF
WFLRQLKERTKTGETVSTKQVETSPVKTKNGIWMLRDSRGSTNMREKRLTV
GCAVTCYRDMGARHARAHSIQIILKDSIPAAKTRVHVAKQFDSKIKPELPVORVH
KGNKRLFSFRKRTYFQ"
Query Match 3.4%; Score 20; DB 5; Length 262395;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 50 ggcctcgtgcctacatgcg 69
|||||
DB 140906 GCGTCTGCTGCTACATGTCG 140925

```

```

RESULT 12
D89210/c D89210 1152 bp mRNA PLN 13-MAR-1998
LOCUS Schizosaccharomyces pombe mRNA, partial cds, clone: SY 1164.
DEFINITION D89210
ACCESSION D89210
VERSION D89210.1 GI:1749627
KEYWORDS
SOURCE Schizosaccharomyces pombe (strain:PR745) cDNA to mRNA,
clone.lib:library of H. Nojima clone:SY1164.
ORGANISM Schizosaccharomyces pombe
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
Schizosaccharomyces.
1 (bases 1 to 1152)
Yoshioaka,S.
Direct Submission
Submitted (15-NOV-1996) to the DDBJ/EMBL/GenBank databases. Sachio
Yoshioaka, Tsukita Cell Axis Project ERATO JST, Kyoto Research Park;
17 Chidouji Minamimachi, Shimokyo-ku, Kyoto, Kyoto 600, Japan
(E-mail: syoshi@cell.tsukita.jst.go.jp, Tel:+81-75-315-7913,
Fax:+81-75-315-6420)
2 (sites)
REFERENCE Yoshioaka,S., Kato,K., Nakai,K., Okayama,H. and Nojima,H.
AUTHORS Identification of Open Reading Frames in Schizosaccharomyces pombe
TITLE
JOURNAL DNA Res. 4 (6), 363-369 (1997)
MEDLINE 98162722
FEATURES
source location/Qualifiers
1..1152
/organism="Schizosaccharomyces pombe"
/strain="PR745"
/db_xref="taxon:4896"
/clone="SY1164"
/clone.lib="library of H. Nojima"
<1..522
/note="similar to Saccharomyces cerevisiae ORF YGL136C,
EMBL Accession Number Z72658"
/codon_start=1
/protein_id="PAA13871.1"
/db_xref="GI:1749628"
/translation="INKEDVYIVYGFAPGSGSYAKKIVGKKGVIGIDIOHAPPG
VLPYIGDIRDPNTITKLEALRLHEPNTNDSIDCRVDAVITSMLEKATGIRIRDA
LSMELCASALHVALTFLKSNCSFTCKRYMGDEADLONLKSHPRFVQVMPKPSALKE
SREAVFYCLEKRP"

```

```

BASE COUNT 325 a 211 c 257 g 359 t
ORIGIN
Query Match 3.2%; Score 19; DB 14; Length 1152;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 233 ctgcaccccccgtttc 251
|||||
DB 853 CTGCATCACCCCTGTTTC 835

```


gene complement(3456. .4076)
/gene="rps3e"
/note="VNG0787G"
CDS complement(3456. .4076)
/gene="rps3e"
/note="Rps3e"
/codon_start=1
/transl_table=11
/product="30S ribosomal protein S3e"
/protein_id="AAG19253.1"
/db_xref="GI:10580361"
/translation="MSERSVSKDQERKMYTVLAPEDRAELGETTAAEEDQYDRT
IOALSDVRDGDNNIKLTFQIDVGSDSASTQVQAELETFDYORSIVRRSSVAVT
VTVTDDYRIQDIPVATYTRKADQSOQHAIKRTIMIDIVEAGEERTFEALNLSIES
RLSAIYDENTYIPLRVEYEKATLEAHPEVEHEEETADPESGE"
gene complement(4073. .4234)
/gene="VNG0788H"
CDS complement(4073. .4234)
/gene="VNG0788H"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng0788h"
/protein_id="AAG19254.1"
/db_xref="GI:10580362"
/translation="MTTVDQDPTVTRIERADTAGMQSTVDDYVNLVTAATVQOT
ANQHTTHOS"
gene complement(4315. .5418)
/gene="VNG0789C"
CDS complement(4315. .5418)
/gene="VNG0789C"
/note="conserved hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng0789c"
/protein_id="AAG19255.1"
/db_xref="GI:10580363"
/translation="MATTGRSSNDSDGVADYAGVSGAEFVRYTAAPTGCVAAGIL
AGYDEVGPVSQVRYVSEPPSHSGAAGSOAAPTTASHVYVYCEAVADASVTVARA
FALAREVGHDPVYTAAGAVADGVEPPADYRAAEVTRPGVGVTTADIAAGLAIS
TLVHASTSGDEAGAFGLAEGLPEPDAAARRLASAVLADATADAPCAADAVHHA
LRPRATPDGMATVGEYADYLVNAYMAAPGVAIGVADRTAVLAMOWHAAPSHA
AVRAASLARHRCVYADVPTAVPVARLLRDYRATPEPAVLAVGTETALATTADANA
VLPSVGTERRAAVTTLDSSDADAVIEFVGGGL"
gene complement(5423. .5890)
/gene="rps15p"
/note="VNG0790G"
CDS complement(5423. .5890)
/gene="rps15p"
/note="Rps15p"
/codon_start=1
/transl_table=11
/product="30S ribosomal protein S15p"
/protein_id="AAG19256.1"
/db_xref="GI:10580364"
/translation="MARMHRRSGSSGSDRPTADEPENSVDDEATIEERYELAFQD
HDSQIGIKRDEQVCTPVDYKLAIGKVTTELLEHDAPELPEDPFRNLLEAVRL
HEVYANGQDHQRALQNTQSKIRLADYIRGKLDIEERYSTETARETIEE"
tRNA 6115. .6186
/gene="trn11"
/product="tRNA-Ala"
6115. .6186
/gene="trn11"
/note="VNG0793G"
gene complement(6477. .8846)
/gene="Htr6"
/note="Htr6"
/codon_start=1
/transl_table=11
/product="Htr6 transducer"
/protein_id="AAG19257.1"

/db_xref="GI:10580365"
/translation="MSDQITGALNMAASVAPDAIRRFQKAFATVTVLVYTAGAG
VFAPQSTETVEHOTTRQIALSNQLEADATGAMMEQOQTHARSISQDBALRDDRRAAP
YTLKLDLLPADVYVSMHVNTEKGTVAESTLAEGLSLADLAPWTTAAVPEEGND
SAVWATEQSYRSPLVNDPEVPAASVPRREGALHVVYTRIQPVDRISDANSYTRRT
ILNTGDEPVLDSNRFDAJLTDLSIAAVRNRTTATTVTSNGRYALSAAPHTDMVY
TSVDTGEAFVADTVGVTVVVLVAVLSIAIVGIALGRHVPLKRLRNADIEEG
TFVDLSTRRIDEVGRILYSPDMRVSLQDRIOBAEAAVEEAMAKAEAEELRTDAD
AQAERAKKATAEASERLOERADYSEWQAVVADGDLTERLDEADBEAMRAVATER
NAMLDGEATTQVAGFPADEVADETLQVATGAEEIETTQVTSERIOELADGALQHD
DLERAGAGMDLSASIOEVAASATVETADAVERAGAGDAESAIDMAETESLS
ADAVDQITLQERMSDIDIEFTTDLAEQNTMALNANIIEAARADKDGCFVAVNE
VKDAEEETKQAAADIESEIQAVOAEETEDVADIRATSEHIDDGVSTVERAAALIEDYV
DAIDANGIOEISDATEDQADATQOSVARYVDVADIOSHYTERADQYSAAEQSNAS
VAETARSDDLRDADALATTVNOFETRADEDEFDADTTVDASADTGD"
gene complement(8843. .9811)
/gene="yufN"
/note="VNG0794G"
CDS complement(8843. .9811)
/gene="yufN"
/note="yufN"
/codon_start=1
/transl_table=11
/product="ABC transporter (lipoprotein)"
/protein_id="AAG19258.1"
/db_xref="GI:10580366"
/translation="MVTALGGLDRSFENDAAANRGAQRARLDHGIATYTNHEPDSIAELG
ETQAVTLAASSNPAYDLVCCVGFQADGLESAAPQOEIVPTADASVASYVE
EOTGSFQAGRLAALLTTRDIDAGGANGPDAKRVGFGVDEPLTGHFGFQGVTH
TASDIDVLAATIGDEVEVEAARTAAAMVADGADIVYHAAGAGVGYLOAAQANGRA
IGVSDQSRSNPRYADVLAISMVRIKIPAVYDALEATIDGGLPDGVDYSLGLATNGVS
VQVCTQICPAIPDVTDALGCTREALVAGDITVTERPQSGGGGV"
BASE COUNT 1353 a 3628 c 3425 g 1722 t
ORIGIN
Query Match 3.2%; Score 19; DB 1; Length 10128;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 386 cgcgcgcacgcagcagcgc 404
|||||
DB 574 CCCCGCCACGACGAGGTC 556
|||||
RESULT 15
AR083116 11873 bp DNA PAT 01-SEP-2000
LOCUS AR083116
DEFINITION Sequence 32 from patent US 5976803.
ACCESSION AR083116
VERSION AR083116.1 GI:10009906
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 11873)
AUTHORS Meek,K.D.
TITLE Genetic test for equine severe combined immunodeficiency disease
JOURNALS Patent: US 5976803-A 32 02-NOV-1999;
FEATURES
SOURCE Location/Qualifiers
1. 11873
/organism="unknown"
BASE COUNT 3585 a 2320 c 2669 g 3299 t
ORIGIN
Query Match 3.2%; Score 19; DB 9; Length 11873;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 211 actctgcgcgtcgtcat 229
|||||
DB 6446 ACTTCTGCCGCTGTCAT 6428

```

RESULT 16
LOCUS AR083115 11878 bp DNA
DEFINITION Sequence 31 from patent US 5976803.
ACCESSION AR083115
VERSION AR083115.1 GI:10009905
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT 3586 a 2323 c 2668 g 3301 t
ORIGIN

Query Match 3.2%; Score 19; DB 9; Length 11878;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 actctgcgcgtcgtcgcgt 229
|||||
Db 6446 ACTTCTGCCGCTGCTGCAT 6428

RESULT 17
LOCUS AR083114 11883 bp DNA
DEFINITION Sequence 28 from patent US 5976803.
ACCESSION AR083114
VERSION AR083114.1 GI:10009904
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT 3588 a 2324 c 2669 g 3302 t
ORIGIN

Query Match 3.2%; Score 19; DB 9; Length 11883;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 actctgcgcgtcgtcgcgt 229
|||||
Db 6449 ACTTCTGCCGCTGCTGCAT 6431

RESULT 18
LOCUS SPBC262 36493 bp DNA
DEFINITION S.pombe chromosome II cosmid c262.
ACCESSION AL022103
VERSION AL022103.1 GI:2956767
KEYWORDS
60s ribosomal protein l16-c; beta-transducin; beta-adaptin; cell
division protein [f1s] homolog; deoxycytidylate deaminase;
glycylpeptide n-tetradecanoyltransferase; histidyl-trna synthetase;
l12lp; manose-6-phosphate isomerase; phosphatidylinositol
phosphate phosphatase; pop1; protein transport protein; rpl16c; sun

```

```

SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
family protein; tetrahydrofolate synthase; transmembrane
transporter; WD repeat protein; WD-40 repeat.
fission yeast.
Schizosaccharomyces pombe
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
Schizosaccharomycetes.
1 (bases 1 to 36493)
Wood, V., Rajandream, M.A., Barrell, B.G., Devlin, K. and Churcher, C.M.
Submitted (04-MAR-1998) European Schizosaccharomycetes genome
sequencing project, Sanger Centre, The Wellcome Trust Genome
Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk
Notes:
Details of yeast sequencing at the Sanger Centre are available on
the World Wide Web.
(URL: http://www.sanger.ac.uk/Projects/S.pombe/)
During 1995 to 1996 about 80% of S. pombe chromosome 1 was
sequenced by the Sanger Centre. The sequencing of the S. pombe
genome is now being continued with funding from The European
Commission. Fourteen European sequencing laboratories, including
the Sanger Centre, are participating in the project.
Protein coding regions (CDS) have been predicted with the help of
computer analysis using the GeneFinder program in Pombase (an ACEDB
database) with additional predictions for the branch-acceptor sites
supplied by the program Sp3splice. CAUTION: It is possible that for
any individual CDS we may have underestimated or overestimated the
number of introns/exons or we may not have chosen the correct
splice donor/acceptor sites.
CDS are numbered using the following system eg SPBC25H2.01c. SP (S.
pombe), B (chromosome 2), c25H2 (cosmid name), .01 (first CDS), c
(complementary strand).
The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous.
The length in codons is given for each CDS.
IMPORTANT: This sequence MAY NOT be the entire insert of the
sequenced clone. It may be shorter because we only sequence
overlapping sections once, or longer, because we arrange for a
small overlap between neighbouring submissions.
Cosmid c262 overlaps 5' with c4B4.
Location/Qualifiers
1. 36493
/organism="Schizosaccharomyces pombe"
/strain="972h-"
/db_xref="taxon:4896"
/chromosome="II"
/clone="cosmid c262"
/map="IIR"
complement(1..1156)
/gene="SPBC262.01c"
/note="l121"
1..1709
/note="8nominal overlap with SPBC4B4 S. pombe chromosome
2"
complement(1..1156)
/partial
/gene="SPBC262.01c"
/function="transmembrane transporter; affects cell cycle
progression"
/note="SPBC262.01c, len:383"
/codon_start=1
/label="l121"
/product="transmembrane transporter l12lp."
/protein_id="CAI17881.1"
/db_xref="GI:2956768"
/db_xref="SPRMBL:043000"
/translation="MALNLRIRAKTSPSPYGIINKVREOKLIIKIDMFLSYCVSYFNY
LDRSSINNAIYSGMODLKNHGELEDIDIVVFPCGIIITQLPESYALQVPRPIMSY
MNLTLGWLMTIFSFVAVSVRALMLIRFPMVAEASTAGTHYILGAWKSEELCKRAGI
FSASGIVGTMFAGYLDQAVHSSLNGSGVMWMLIDGIIITPLSLGLFEPDVP
ENTKAVYFEQKEKELSEFKRLPAPKKRPLTIAIKDIVSMRYIGICLMTISGETQA
IANNVAGMGOMKSNKSESVAOINNNPTVITAVGVSTLGSVYSIDSLACNPMPFCLF
LCVITIVSATILLANVDPGAKFAFAFGAGCTYAGAAVAFSANDICRND"
FEATURES
source
gene
misc_feature
CDS

```

```
gene
2337..5567
/gene="SPBC262.02"
CDS
2337..5567
/note="SPBC262.02"
/note="SPBC262.02, len:1075, SIMILARITY:Saccharomyces cerevisiae, Q12271, YOR109W, phosphatidylinositol phosphate phosphatase (1107 aa), fasta scores: opt: 1630, E():0, (39.0% identity in 1100 aa)"
/codon_start=1
/label="SPBC262.02"
/product="phosphatidylinositol phosphate phosphatase; synaptojanin-like"
/protein_id="CAI17882.1"
/db_xref="GI:2956769"
/db_xref="SPTREMBL:O43001"
/translation="MOCLLRKPRSLAVNKDHALMFHSVPQNKNSLVCVAEFTALS EKPLEGFRKISSHRVIGTLGLEEGSNFLCIVGASEVAVRDKERERFIMECFVS VNRSMNPHIROENYSPIDPGYDPTQGVDSYKAAEPFSLRKLITNGSYFSLDFD ITTRLOJRTSOTMTPEQYDSMHTQFMNNEFMRLOLIERSHLNGDEKSDLGCRFFTC AIRGEASTEOFKLGIOTIRLSLISRLSRLGRFLSGVDGDNVANVFETELIIS SKVGVSCVGVVLRSGOIVGTALMIFCKESCILPKNEGVTKKGLGCVSGKNAV AIRPDYDTGLCTTSHLAGYTYNDESDHDTRTIASGLRRGRSTFENHDYVWFGD FNRISLTVLEHVPCIAQGLKSLYFEYDQUNKOMLTGKVPFESLPTPEPTFKDI GTDIYDSLKHVPAMTDRLIRGELVPHYSOYPLYSDBRPYATYEAIVAVRE KKKLFEELYNQKROEVDAQSOTYTLIDINGSAGPNLPHLPANGVDIKQPSSE RSKWPFDDGLPAKSIAPGPPEYRLNPSRPINPEPAPEPMISNTQSPFKKSLID SIPALSPASSIASVSSORSSTIIPKPKNTKRDHYIAPVKKLLPPRSSSSS GVPAPNTPVNVPPPPPKSSASQSRGDLASSPEBSISWKLV"
complement(join(3906..6005,6056..6187,6252..6313, 6364..6378))
/gene="SPBC262.03c"
complement(join(5906..6005,6056..6187,6252..6313, 6364..6378))
/gene="SPBC262.03c"
/note="SPBC262.03c, len:101, SIMILARITY:Saccharomyces cerevisiae, S6B2_YEAST, protein transport protein sec61 beta 2 subunit., (88 aa), fasta scores: opt: 242, E():2.1e-09, (46.3% identity in 82 aa)"
/codon_start=1
/label="SPBC262.03c"
/product="putative protein transport protein: yeast sec61 beta 2 subunit-like"
/protein_id="CAI17883.1"
/db_xref="GI:2956770"
/db_xref="SPTREMBL:O43002"
/translation="MSSSTKASGVKNSANASAPGPKSOIRRAAVEKNTKSNCPAG ARAGAGPSTPLTKLYTDEASGFKVPVVMVVLGVFIASVFLHIVARILKRFAS"

misc_feature
complement(6006..6020)
/note="ctaactgaataag, splice branch and acceptor"
misc_feature
complement(6050..6055)
/note="gtaagc, splice donor sequence"
misc_feature
complement(6188..6210)
/note="ctaactttgtgtgactgtag, splice branch and acceptor"
misc_feature
complement(6246..6251)
/note="gttcgt, splice donor sequence"
misc_feature
complement(6314..6324)
/note="ctaactgaatg, splice branch and acceptor"
misc_feature
complement(6358..6363)
/note="gtaagt, splice donor sequence"
misc_feature
complement(6975..7463)
/gene="SPBC262.04c"
complement(6975..7463)
/gene="SPBC262.04c"
/note="SPBC262.04c, len:161, SIMILARITY:Saccharomyces cerevisiae, YIF1_YEAST, required for maintenance of
```

```
mitochondrial DNA, (145 aa), fasta scores: opt: 357, E():7.5e-17, (44.0% identity in 141 aa)
sequence possibly needs truncating to second met"
/codon_start=1
/label="SPBC262.04c"
/product="hypothetical protein"
/protein_id="CAI17884.1"
/db_xref="GI:2956771"
/db_xref="SWISS-PROT:O43003"
/translation="MLRALGSRLLVAPRAVRSFQSLRPPRFTHKMSFTKPLNS PKLSAGPVNOAIKANGVYCSGOIPVANGVIGTGVDDYRQCLINIOEVLTEAGSS LNKIVKVIIFLADMDDFAAVNVKYTEVLPDKPARSCVAVTVPLSGVYIEICIA LE"
join(8051..8091,8150..8702)
/gene="SPBC262.05"
/note="rpl16c"
join(8051..8091,8150..8702)
/gene="SPBC262.05"
/note="SPBC262.05, len:196"
/codon_start=1
/label="rpl16c"
/product="60s ribosomal protein l16-c."
/protein_id="CAI17885.1"
/db_xref="GI:2956772"
/db_xref="SWISS-PROT:O43004"
/translation="MSEPRKVIITIDAGHLMGRVLAQVYAKOLAGOKVYVRCHEMT SGHFRNKLALALRAKCRNPGRGAFHFRAPRITTKVAGRLPKHTTGTALKN LQALEGLPPEPKOKRLVPAALVLRKPSKCTTGIRLSSEVGMKYNKIVLSLEER RKISAEVYQAKSANKHINAKTKSSVNEKLAFCY"
8092..8097
/note="gta, splice donor sequence"
8135..8149
/note="ctaactgacttag, splice branch and acceptor"
8361..8432
/gene="SPBC262.05"
/note="PS00783 Ribosomal protein L13 signature"
complement(join(8859..10245,10290..10408,10453..10698, 10752..10797,10835..11043,11090..11116))
/gene="SPBC262.06c"
complement(join(8859..10245,10290..10408,10453..10698, 10752..10797,10835..11043,11090..11116))
/gene="SPBC262.06c"
/note="SPBC262.06c, len:676, SIMILARITY:Drosophila melanogaster, Q24253, beta-adaptin drosophila 1., (921 aa), fasta scores: opt: 1724, E():0, (45.2% identity in

Query Match 3.2%; Score 19; DB 15; Length 36493;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 233 ctgcatcacccctgtttc 251
Db 27668 CTGCATCACCCCTGTTC 27686

RESULT 19
AC068903/c 82033 bp DNA 20-SEP-2000
LOCUS DEFINITION Mus musculus chromosome 2 clone ct7-196d19 strain 129/Sv ES cell
AC068903 line c77, complete sequence.
ACCESSION AC068903
VERSION AC068903.11 GI:9837949
KEYWORDS HMG.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 82033)
AUTHORS Shaull,S., Rahhal,R., Yao,Z., Rithidech,K. and Roe,B.A.
JOURNAL Mus musculus Chromosome 2 BAC Clone ct7b-196d19
REFERENCE 2 (bases 1 to 82033)
AUTHORS Shaull,S., Rahhal,R., Yao,Z., Rithidech,K. and Roe,B.A.
```

TITLE Direct Submission
JOURNAL Submitted (11-MAY-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 3 (bases 1 to 82033)
AUTHORS Shaull,S., Rahhal,R., Yao,Z., Rithidech,K. and Roe,B.A.
JOURNAL Direct Submission
Submitted (20-SEP-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Aug 17, 2000 this sequence version replaced g1:9797790.
FEATURES
source location/Qualifiers
1..82033
/organism="Mus musculus"
/strain="129/Sv ES cell line Cj7"
/db_xref="taxon:10090"
/chromosome="2"
/clone="ct7-196d19"
/clone_1lb="Caltech CITB-BAC library"
BASE COUNT 23822 a 16503 c 15652 g 26056 t
ORIGIN

Query Match 3.2%; Score 19; DB 94; Length 82033;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaaccacagagagaagaa 555
|||||
DB 61880 GAACCAACAGAGAGAA 61862

RESULT 20
AC068952/c AC068952 95292 bp DNA ROD 18-AUG-2000
LOCUS Mus musculus chromosome 2 clone ct7-254f1 strain 129/Sv ES cell
DEFINITION line Cj7, complete sequence.
ACCESSION AC068952
VERSION AC068952.10 GI:9845112
KEYWORDS HTG.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 95292)
AUTHORS Shaull,S., Rahhal,R., Yao,Z., Rithidech,K. and Roe,B.A.
JOURNAL Mus musculus Chromosome 2 BAC Clone cltb-254f1
TITLE Unpublished
AUTHORS 2 (bases 1 to 95292)
JOURNAL Shaull,S., Rahhal,R., Yao,Z., Rithidech,K. and Roe,B.A.
TITLE Direct Submission
Submitted (13-MAY-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 3 (bases 1 to 95292)
AUTHORS Shaull,S., Rahhal,R., Yao,Z., Rithidech,K. and Roe,B.A.
JOURNAL Direct Submission
Submitted (18-AUG-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Aug 18, 2000 this sequence version replaced g1:9795634.
FEATURES
source location/Qualifiers
1..95292
/organism="Mus musculus"
/strain="129/Sv ES cell line Cj7"
/db_xref="taxon:10090"
/chromosome="2"
/clone="ct7-254f1"
/clone_1lb="Caltech CITB-BAC library"
BASE COUNT 28084 a 18825 c 18104 g 30279 t
ORIGIN

Query Match 3.2%; Score 19; DB 94; Length 95292;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaaccacagagagaagaa 555
|||||
DB 76659 GAACCAACAGAGAGAA 76641

RESULT 21
AL354832.1
WPCOMMENT
Sequence split into 4 fragments LOCUS AL354832 Accession AL354832
Fragment Name Begin End
AL354832_0 1 110000
AL354832.1 100001 210000
AL354832.2 200001 310000
AL354832_3 300001 372955
Continuation (2 of 4) of AL354832 from base 100001 (AL354832 Homo sapiens chromosome

Query Match 3.2%; Score 19; DB 79; Length 110000;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 ctgccgtgctgcatcccc 233
|||||
DB 66912 CTGCCGCTGCTGCATCCCC 66930

RESULT 22
AC024591/c AC024591 111071 bp DNA HTG 18-JUL-2000
LOCUS Homo sapiens chromosome 16 clone RP11-511G21, WORKING DRAFT
DEFINITION SEQUENCE, 13 ordered pieces.
ACCESSION AC024591
VERSION AC024591.3 GI:9256450
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 111071)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 111071)
JOURNAL DOE Joint Genome Institute.
TITLE Direct Submission
Submitted (29-FEB-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 18, 2000 this sequence version replaced g1:7705016.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 602754
Center clone name: RPC1-11_511G21

Summary Statistics
Consensus quality: 104021 bases at least Q40
Consensus quality: 108348 bases at least Q30
Consensus quality: 109319 bases at least Q20
Estimated insert size: 118930; agarose-fp estimation
Estimated insert size: 110521; sum-of-configs estimation
Quality coverage: 8.44 in Q20 bases; agarose-fp estimation
Quality coverage: 9.08 in Q20 bases; sum-of-configs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes

| | |
|---|--|
| * | of the gps between them are based on estimates that have |
| * | provided by the submittor. |
| * | This sequence will be replaced |
| * | by the finished sequence as soon as it is available and |
| * | the accession number will be preserved. |
| * | 1 3999: contig of 3999 bp in length |
| * | 4000 4099: gap of unknown length |
| * | 4100 6224: contig of 2125 bp in length |
| * | 6225 6324: gap of unknown length |
| * | 6325 14602: contig of 8278 bp in length |
| * | 14603 14702: gap of unknown length |
| * | 14703 54245: contig of 39543 bp in length |
| * | 54246 54345: gap of unknown length |
| * | 54346 58509: contig of 4224 bp in length |
| * | 58570 58669: gap of unknown length |
| * | 58670 59480: contig of 811 bp in length |
| * | 59481 59580: gap of unknown length |
| * | 59581 60683: contig of 1103 bp in length |
| * | 60684 60783: gap of unknown length |
| * | 60784 62950: contig of 2167 bp in length |
| * | 62951 63051: gap of unknown length |
| * | 63051 64092: contig of 1042 bp in length |
| * | 64093 64193: gap of unknown length |
| * | 64193 85067: contig of 20875 bp in length |
| * | 85068 85167: gap of unknown length |
| * | 85168 91149: contig of 5962 bp in length |
| * | 91150 91249: gap of unknown length |
| * | 91250 108336: contig of 16867 bp in length |
| * | 108237 108336: gap of unknown length |
| * | 108337 11071: contig of 2735 bp in length. |

| FEATURES | location | unaffiliers |
|------------|--------------------------|-------------------------------------|
| SOURCE | 1. | 111071 |
| | /organism="Homo sapiens" | |
| | /db_xref="taxon:9606" | |
| | /chromosome="16" | |
| | /clone="RP11-511c21" | |
| BASE COUNT | /clone_id="RC1 | human BAC library 11" |
| ORIGIN | 27517 a | 27152 c 26256 g 26944 t 1202 others |

Query Match 3.2%; Score 19; DB 68; Length 111071;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      280 gccgcgycgcgatcgcgtg    298  
          |||||  
Db 19460 GCCGCGCGCGCATGGCTG 19442
```

| RESULT | 23 |
|------------|---|
| AC020962 | |
| LOCUS | AC020962 128562 bp DNA |
| DEFINITION | Mus musculus clone RP21-126H5, WORKING DRAFT SEQUENCE, 21 unordered |
| ACCESSION | AC020962 |
| VERSION | AC020962.1 GI:6691268 |
| KEYWORDS | HTG; HTGS_PHASE1; HTGS_DRAFT. |
| SOURCE | house mouse. |
| ORGANISM | Mus musculus |

| | |
|-----------|--|
| REFERENCE | 1 (bases 1 to 128562) |
| AUTHORS | DOE Joint Genome Institute. |
| TITLE | Sequencing of Mouse |
| JOURNAL | Unpublished |
| REFERENCE | 2 (bases 1 to 128562) |
| AUTHORS | DOE Joint Genome Institute. |
| TITLE | Direct Submision |
| JOURNAL | Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA |
| COMMENT | -----Genome Center Center: Joint Genome Institute |

Center Code: JGI
Web site: <http://www.jgi.doe.gov>
-----Summary Statistics

| | | |
|------------------------|--------------|-----------------------------------|
| Consensus quality: | 11351.4 | bases at least Q40 |
| Consensus quality: | 11105.5 | bases at least Q30 |
| Consensus quality: | 11628.5 | bases at least Q20 |
| Estimated insert size: | 12856.2 | sum-of-conflics estimation |
| Estimated insert size: | 12200.1 | pulse field gel estimation |
| Quality coverage: | 7.89x in Q20 | bases; pulse field gel estimation |
| Quality coverage: | 7.48x in Q20 | bases; sum-of-conflics estimation |

* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

| | | |
|------------|---------------------------------|---------------------------------------|
| 1 | 1198: | contig of 1198 bp in length |
| * | | gap of unknown length |
| * | 1199 | 2531: contig of 1333 bp in length |
| * | | gap of unknown length |
| * | 2532 | 4140: contig of 1609 bp in length |
| * | | gap of unknown length |
| * | 4141 | 5328: contig of 1188 bp in length |
| * | | gap of unknown length |
| * | 5329 | 6673: contig of 1345 bp in length |
| * | | gap of unknown length |
| * | 6674 | 8037: contig of 1364 bp in length |
| * | | gap of unknown length |
| * | 8038 | 9583: contig of 1546 bp in length |
| * | | gap of unknown length |
| * | 9584 | 11046: contig of 1463 bp in length |
| * | | gap of unknown length |
| * | 11047 | 12841: contig of 1795 bp in length |
| * | | gap of unknown length |
| * | 12842 | 14056: contig of 1215 bp in length |
| * | | gap of unknown length |
| * | 14057 | 15421: contig of 1365 bp in length |
| * | | gap of unknown length |
| * | 15422 | 16973: contig of 1552 bp in length |
| * | | gap of unknown length |
| * | 16974 | 18746: contig of 1773 bp in length |
| * | | gap of unknown length |
| * | 18747 | 20176: contig of 1430 bp in length |
| * | | gap of unknown length |
| * | 20177 | 21643: contig of 1467 bp in length |
| * | | gap of unknown length |
| * | 21644 | 22955: contig of 1312 bp in length |
| * | | gap of unknown length |
| * | 22956 | 24395: contig of 1440 bp in length |
| * | | gap of unknown length |
| * | 24396 | 26493: contig of 2098 bp in length |
| * | | gap of unknown length |
| * | 26494 | 46982: contig of 20489 bp in length |
| * | | gap of unknown length |
| * | 46983 | 78189: contig of 31207 bp in length |
| * | | gap of unknown length |
| * | 78190 | 128562: contig of 50373 bp in length. |
| FEATURES | | |
| SOURCE | | |
| | 1. 128562 | |
| | /organism="Mus musculus" | |
| | /db_xref="taxon:10090" | |
| | /clone="RP21-126H5" | |
| BASE COUNT | 31926 a 32661 c 31576 g 32269 t | 130 others |
| ORIGIN | | |

| | | | | |
|-----------------------|-----------------|---------------|-----------|----------------|
| Query Match | 3.2%; | Score 19; | DB 66; | Length 128562; |
| Best Local Similarity | 100.0%; | Pred. No. 28; | | |
| Matches 19; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |

QY 280 gccgcggcgatcgctg 298
 Db 117029 GCCGCGCGGATCGCTG 117047

RESULT 24
 LOCUS AC073882 130734 bp DNA ROD 24-AUG-2000
 DEFINITION Mus musculus chromosome 2 clone ct7-305f12 strain 129/SVES Cell
 ACCESSION AC073882
 KEYWORDS AC073882.4 GI:9690337
 SOURCE HTG.
 ORGANISM house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 130734)
 JOURNAL Unpublished
 TITLE 2 (bases 1 to 130734)
 AUTHORS Jia, H., Lin, S., Rithideck, K. and Roe, B.A.
 JOURNAL Submitted (01-JUL-2000) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 REFERENCE 3 (bases 1 to 130734)
 AUTHORS Jia, H., Lin, S., Rithideck, K. and Roe, B.A.
 JOURNAL Direct Submission
 TITLE Submitted (04-AUG-2000) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 REFERENCE 4 (bases 1 to 130734)
 AUTHORS Jia, H., Lin, S., Rithideck, K. and Roe, B.A.
 JOURNAL Direct Submission
 TITLE Submitted (24-AUG-2000) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 COMMENT On Aug 4, 2000 this sequence version replaced g1:9280733.
 FEATURES
 source
 1. .130734
 /organism="Mus musculus"
 /strain="129/SVES Cell Line Cj7"
 /db_xref="taxon:10090"
 /chromosome="2"
 /clone="ct7-305f12"
 /clone.lib="Caltech CITB-BAC library"
 BASE COUNT 41025 a 25364 c 26653 g 37692 t

Query Match 3.2%; Score 19; DB 94; Length 130734;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaacacacagagaaga 555
 Db 65070 GAAACCAACAGAGAGAA 65088

RESULT 25
 LOCUS HSJ1057D4 131888 bp DNA PRI 01-MAR-2001
 DEFINITION Human DNA sequence from clone RP5-1057D4 on chromosome 20 Contains
 a spermidine synthase (SPDSY) pseudogene, a CpG island, ESTs, STS
 and GSSs, complete sequence.
 ACCESSION AL121777
 VERSION AL121777.39 GI:10862842
 KEYWORDS HTG; CpG island; spermidine synthase.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 131888)
 AUTHORS Blakey, S.
 JOURNAL Direct Submission
 TITLE Submitted (26-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerrysanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Oct 17, 2000 this sequence version replaced g1:10716550.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
 on the WORMPEP database can be found at
 http://www.sanger.ac.uk/projects/c_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 20, constructed by the Sanger Centre Chromosome 20
 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr20
 This sequence is the entire insert of clone RP5-1057D4 The true
 left end of clone RP1-73E16 is at 71240 in this sequence. The true
 right end of clone CTD-2653D5 is at 84358 in this sequence. This
 sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30): an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. RP5-1057D4 is from
 the library RPC1-5 constructed by the group of Pieter de Jong. For
 further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pCYPAC2.
 FEATURES
 source
 Location/Qualifiers
 1. .131888
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="20"
 /clone="RP5-1057D4"
 /clone.lib="RPC1-5"
 repeat_region
 313. .930
 /note="12 repeat: matches 2043. .2749 of consensus"
 1411. .1543
 /note="MIR repeat: matches 80. .222 of consensus"
 1623. .1934
 /note="AluY repeat: matches 1. .311 of consensus"
 2162. .2215
 /note="12 repeat: matches 2696. .2749 of consensus"
 2444. .2646
 /note="12 repeat: matches 2353. .2521 of consensus"
 2929. .3407
 /note="match: GSS: Em: A0520865"
 3104. .3214
 /note="12 repeat: matches 2629. .2746 of consensus"
 3638. .3839
 /note="MIR repeat: matches 32. .256 of consensus"
 3916. .4051
 /note="AluX repeat: matches 1. .123 of consensus"
 4052. .4341
 /note="AluY repeat: matches 1. .295 of consensus"
 4342. .4477
 /note="AluX repeat: matches 123. .298 of consensus"
 5436. .5600
 /note="AluX repeat: matches 170. .299 of consensus"
 5601. .5892
 /note="AluY repeat: matches 1. .291 of consensus"
 5893. .6050
 /note="AluX repeat: matches 1. .170 of consensus"
 6179. .6230

```

repeat_region /note="MR repeat: matches 206. .256 of consensus
6186. .6565
/note="L2 repeat: matches 2292. .2731 of consensus"
6567. .6606
repeat_region /note="L2 repeat: matches 2148. .2189 of consensus"
6607. .6913
repeat_region /note="AluJo repeat: matches 1. .307 of consensus"
6914. .7005
repeat_region /note="L2 repeat: matches 2064. .2148 of consensus"
7104. .7559
repeat_region /note="MTR1G repeat: matches 31. .512 of consensus"
8768. .8969
repeat_region /note="MER20 repeat: matches 1. .216 of consensus"
9696. .10011
repeat_region /note="L2 repeat: matches 2357. .2734 of consensus"
11637. .12065
repeat_region /note="MTR1H repeat: matches 20. .531 of consensus"
12302. .12377
repeat_region /note="MR repeat: matches 115. .190 of consensus"
13520. .13783
repeat_region /note="L2 repeat: matches 2432. .2749 of consensus"
13953. .14155
repeat_region /note="MTR1J repeat: matches 267. .468 of consensus"
14420. .14705
repeat_region /note="AluDb repeat: matches 1. .294 of consensus"
complement(14545. .15063)
misc_feature /note="match: GSS: Em:A0713490"
complement(14589. .15060)
misc_feature /note="match: GSS: Em:A0760291"
complement(14602. .15072)
misc_feature /note="match: GSS: Em:A0885065"
15070. .15576
misc_feature /note="match: GSS: Em:A0709181"
15088. .15403
misc_feature /note="match: GSS: Em:A0198342"
15088. .15583
misc_feature /note="match: GSS: Em:A0680811"
15094. .15492
misc_feature /note="match: GSS: Em:A0082327"
15423. .15656
repeat_region /note="L2 repeat: matches 1609. .1854 of consensus"
15716. .15894
repeat_region /note="MER20 repeat: matches 4. .187 of consensus"
15898. .16135
repeat_region /note="L2 repeat: matches 1915. .2159 of consensus"
16156. .16457
repeat_region /note="AluXs repeat: matches 3. .304 of consensus"
16458. .16568
repeat_region /note="L2 repeat: matches 2159. .2263 of consensus"
16569. .16869
repeat_region /note="AluXs repeat: matches 1. .301 of consensus"
16870. .17097
repeat_region /note="L2 repeat: matches 2263. .2531 of consensus"
17103. .17406
repeat_region /note="AluXs repeat: matches 1. .301 of consensus"
18176. .18380
repeat_region /note="MR repeat: matches 36. .260 of consensus"
19512. .19745
repeat_region /note="MR repeat: matches 13. .261 of consensus"
complement(20625. .21280)
misc_feature /note="match: GSS: Em:A0383848"
21138. .23316
misc_feature /note="CpG island"
evidence=not_experimental
repeat_region 21792. .22251
/note="230 copies 2 mer cc 55% conserved"
21814. .22282
/note="Single clone region. Assembly confirmed by
restriction digest."
21829. .22122
/note="6 copies 49 mer 64% conserved"
21832. .22251
/note="12 copies 35 mer 57% conserved"

```

| Query Match | Best Local Similarity | Score 19: | DB 93: | Length 131888: |
|--|-----------------------|-----------|--------|----------------|
| Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0 | | | | |
| 215 ctgcgcgtctgcatccccc 233 | | | | |
| Db 22024 CTGCCGCTGCTCATCCCC 22042 | | | | |
| RESULT 26 | | | | |
| AC084179/c | | | | |
| LOCUS | | | | |
| DEFINITION | | | | |
| AC084179 | | | | |
| VERSION | | | | |
| KEYWORDS | | | | |
| SOURCE | | | | |
| ORGANISM | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |

TITLE
JOURNAL
COMMENT

Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
Dearrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,
Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Glade, S., Goyette, M.,
Graham, L., Grand-pierre, N., Hagos, B., Heatford, A., Horton, L.,
Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Larocque, K.,
Lamard, R., Landers, T., Lenock, J., Levine, R., Lieu, C., Liu, G.,
Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K.,
McPheters, R., Melidim, J., Meneus, L., Mihova, T., Mienga, V.,
Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,
O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K.,
Pierre, N., Pisan, C., Pollara, V., Raymond, C., Rieback, M., Riley, R.,
Rogov, P., Rotman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,
Sougnaz, C., Spencer, B., Stange, Thoman, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testafay, S., Theodore, J.,
Turrell, A., Travers, M., Triglio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
Zimmer, A. and Zody, M.

Direct Submission
Submitted (14-OCT-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 15, 2001 this sequence version replaced g1:10801407.
All repeats were identified using RepeatMasker:
Sift, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute / MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center project name: L11428
Center clone name: L148_P_16

Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator; Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 13223 bases at least Q40
Consensus quality: 139576 bases at least Q30
Consensus quality: 142136 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 143209; sum-of-ctnigs
Quality coverage: 3.0 in Q20 bases; sum-of-ctnigs
Quality coverage: 3.6 in Q20 bases; sum-of-ctnigs

NOTE: This is a 'working draft' sequence. It currently
consists of 39 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1353: contig of 1353 bp in length
1354 1453: gap of 100 bp
1454 2001: contig of 548 bp in length
2002 2101: gap of 100 bp
2102 3124: contig of 1023 bp in length
3125 3224: gap of 100 bp
3225 4546: contig of 1322 bp in length
4547 4646: gap of 100 bp
4647 6015: contig of 1369 bp in length
6016 6115: gap of 100 bp
6116 7299: contig of 1184 bp in length
7300 7399: gap of 100 bp
7400 8836: contig of 1437 bp in length
8837 8936: gap of 100 bp
8937 10282: contig of 1346 bp in length
10283 10382: gap of 100 bp
10383 11352: contig of 970 bp in length
11353 11452: gap of 100 bp
11453 12856: contig of 1404 bp in length
12857 12956: gap of 100 bp

FEATURES
source

12957 14187: contig of 1231 bp in length
14188 14287: gap of 100 bp
14288 15977: contig of 1690 bp in length
15978 16077: gap of 100 bp
16078 17844: contig of 1767 bp in length
17845 17944: gap of 100 bp
17945 19861: contig of 1917 bp in length
19862 19961: gap of 100 bp
19962 22368: contig of 2407 bp in length
22369 22468: gap of 100 bp
22469 24485: contig of 2017 bp in length
24486 24585: gap of 100 bp
24586 26722: contig of 2137 bp in length
26723 26822: gap of 100 bp
26823 28983: contig of 2161 bp in length
28984 29083: gap of 100 bp
29084 31858: contig of 2775 bp in length
31859 31958: gap of 100 bp
31959 33783: contig of 1825 bp in length
33784 33883: gap of 100 bp
33884 36259: contig of 2376 bp in length
36260 36359: gap of 100 bp
36360 39396: contig of 3037 bp in length
39397 39496: gap of 100 bp
39497 43329: contig of 3833 bp in length
43330 43429: gap of 100 bp
43430 46057: contig of 2628 bp in length
46058 46157: gap of 100 bp
46158 50066: contig of 3909 bp in length
50067 50166: gap of 100 bp
50167 53727: contig of 3561 bp in length
53728 53827: gap of 100 bp
53828 57631: contig of 3804 bp in length
57632 57731: gap of 100 bp
57732 61304: contig of 3573 bp in length
61305 61404: gap of 100 bp
61405 63868: contig of 2464 bp in length
63869 63968: gap of 100 bp
63969 68393: contig of 4425 bp in length
68394 68493: gap of 100 bp
68494 72466: contig of 3973 bp in length
72467 72566: gap of 100 bp
72567 77027: contig of 4461 bp in length
77028 77127: gap of 100 bp
77128 81524: contig of 4397 bp in length
81525 81624: gap of 100 bp
81625 88873: contig of 7249 bp in length
88874 88973: gap of 100 bp
88974 96287: contig of 7314 bp in length
96288 96387: gap of 100 bp
96388 117899: contig of 2152 bp in length
117900 117999: gap of 100 bp
118000 123663: contig of 5664 bp in length
123664 123763: gap of 100 bp
123764 137598: contig of 13835 bp in length
137599 137698: gap of 100 bp
137699 147009: contig of 9311 bp in length.
Location/Qualifiers
1. 147009
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5"
/clone="RP11-148P16"
/clone_1lb="RPC1-11 Human Male BAC"
1. 1353
/note="assembly_fragment"
clone_end:sp6
vector_side:left
1454. .2001
/note="assembly_fragment"
2102. 3124
/note="assembly_fragment"

misc_feature

1. 1353
/note="assembly_fragment"

misc_feature

1454. .2001
/note="assembly_fragment"
2102. 3124
/note="assembly_fragment"


```

misc_feature 3225..4546
                /note="assembly-fragment"
misc_feature 4647..6015
                /note="assembly-fragment"
misc_feature 6116..7299
                /note="assembly-fragment"
misc_feature 7400..8836
                /note="assembly-fragment"
misc_feature 8937..10282
                /note="assembly-fragment"
misc_feature 10383..11352
                /note="assembly-fragment"
misc_feature 11453..12856
                /note="assembly-fragment"
misc_feature 12957..14187
                /note="assembly-fragment"
misc_feature 14288..15977
                /note="assembly-fragment"
misc_feature 16078..17844
                /note="assembly-fragment"
misc_feature 17945..19861
                /note="assembly-fragment"
misc_feature 19962..22368
                /note="assembly-fragment"
misc_feature 22469..24485
                /note="assembly-fragment"
misc_feature 24586..26722
                /note="assembly-fragment"
misc_feature 26823..28983
                /note="assembly-fragment"
misc_feature 29084..31858
                /note="assembly-fragment"
misc_feature 31959..33783
                /note="assembly-fragment"
misc_feature 33884..36259
                /note="assembly-fragment"
misc_feature 36360..39396
                /note="assembly-fragment"

```

```

Query Match      3.2%  Score 19:  DB 76;  Length 147009;
Best Local Similarity 100.0%;  Pred. No. 27;
Matches 19;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

Oy 190 agtagagtcataccaata 208
|||||
DB 10253 AGTAGAGTCATCAATCAATA 10235

```

```

RESULT 27
AC083840 156599 bp DNA HTG 03-MAR-2001
LOCUS Homo sapiens chromosome 8 clone RP11-7319 map 8, WORKING DRAFT
DEFINITION
SEQUENCE, 37 unordered pieces.
AC083840
AC083840.2 GI:13184085
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 156599)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 8, clone RP11-7319
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 156599)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L.,
Bouhageltier,B., Brown,A., Burkett,G., Campoliano,A., Castle,A.,
Chapel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRoque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,

```

TITLE JOURNAL COMMENT

Macdonald, P., Margulis, N., McCarthy, M., McEwan, P., McKernan, R., McPheeters, R., Meldrum, J., Meneus, L., Mhova, T., Mlenga, Y., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivier, T.M., Oliver, J., Peterson, K., Pierre, N., Pisanu, C., Pollara, V., Raymond, C., Rieback, M., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Sougnuez, C., Spencer, B., Stange-Thumann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirelli, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (03-OCT-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 2, 2001 this sequence version replaced gi:10518398.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: 111308
Center clone name: 73.1.9

Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 141695 bases at least Q40
Consensus quality: 148947 bases at least Q30
Consensus quality: 151617 bases at least Q20
Insert size: 152899; sum-of-contigs
Quality coverage: 4.0 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 37 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1      816: contig of 816 bp in length
*
* 917 916: gap of 100 bp
*
* 917 2094: contig of 1178 bp in length
*
* 2095 2194: gap of 100 bp
*
* 2195 3169: contig of 975 bp in length
*
* 3170 3269: gap of 100 bp
*
* 3270 3766: contig of 497 bp in length
*
* 3767 3866: gap of 100 bp
*
* 3867 5257: contig of 1391 bp in length
*
* 5258 5357: gap of 100 bp
*
* 5358 7176: contig of 1819 bp in length
*
* 7177 7276: gap of 100 bp
*
* 7277 8366: contig of 1090 bp in length
*
* 8367 8466: gap of 100 bp
*
* 8467 9770: contig of 1304 bp in length
*
* 9771 9870: gap of 100 bp
*
* 9871 11023: contig of 1153 bp in length
*
* 11024 11123: gap of 100 bp
*
* 11124 12210: contig of 1087 bp in length
*
* 12211 12310: gap of 100 bp
*
* 12311 13716: contig of 1406 bp in length
*
* 13717 13816: gap of 100 bp
*
* 13817 14696: contig of 880 bp in length
*
* 14697 14796: gap of 100 bp
*
* 14797 16050: contig of 1254 bp in length
*
* 16051 16150: gap of 100 bp
*
* 16151 18691: contig of 2541 bp in length
*
* 18692 18791: gap of 100 bp
*
* 18792 20474: contig of 1683 bp in length

```

* 20475 20574: gap of 100 bp
* 20575 22564: contig of 1990 bp in length
* 22565 22664: gap of 100 bp
* 22665 25349: contig of 2685 bp in length
* 25350 25449: gap of 100 bp
* 25450 28731: contig of 3282 bp in length
* 28732 28831: gap of 100 bp
* 28832 41793: contig of 12962 bp in length
* 41794 41893: gap of 100 bp
* 41894 45562: contig of 3669 bp in length
* 45563 45662: gap of 100 bp
* 45663 49245: contig of 3583 bp in length
* 49246 49345: gap of 100 bp
* 49346 52701: contig of 3356 bp in length
* 52702 52801: gap of 100 bp
* 52802 57587: contig of 4786 bp in length
* 57588 57687: gap of 100 bp
* 57688 60742: contig of 3055 bp in length
* 60743 60842: gap of 100 bp
* 60843 64655: contig of 3813 bp in length
* 64656 64755: gap of 100 bp
* 64756 68878: contig of 4123 bp in length
* 68879 68978: gap of 100 bp
* 68979 73540: contig of 4562 bp in length
* 73541 73640: gap of 100 bp
* 73641 79164: contig of 5524 bp in length
* 79165 79264: gap of 100 bp
* 79265 87667: contig of 8403 bp in length
* 87668 87767: gap of 100 bp
* 87768 96110: contig of 8343 bp in length
* 96111 96210: gap of 100 bp
* 96211 104171: contig of 7961 bp in length
* 104172 104271: gap of 100 bp
* 104272 110622: contig of 6351 bp in length
* 110623 110722: gap of 100 bp
* 110723 121728: contig of 11006 bp in length
* 121729 121828: gap of 100 bp
* 121829 129949: contig of 8121 bp in length
* 129950 130049: gap of 100 bp
* 130050 139747: contig of 9698 bp in length
* 139748 139847: gap of 100 bp
* 139848 153810: contig of 13963 bp in length
* 153811 153910: gap of 100 bp
* 153911 156599: contig of 2689 bp in length.
Location/Qualifiers
1. 156599

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-7319"
/clone_1lb="RPC1-11 Human Male BAC"
1. 816

misc_feature
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
917. 2094
/note="assembly_fragment"
2195. 3169
/note="assembly_fragment"
3270. 3766
/note="assembly_fragment"
3867. 5257
/note="assembly_fragment"
5358. 7176
/note="assembly_fragment"
7277. 8366
/note="assembly_fragment"
8467. 9770
/note="assembly_fragment"
9871. 11023
/note="assembly_fragment"
11124. 12210

misc_feature
/note="assembly_fragment"
12311. 13716
/note="assembly_fragment"
13817. 14696
/note="assembly_fragment"
14797. 16050
/note="assembly_fragment"
16151. 18691
/note="assembly_fragment"
18792. 20474
/note="assembly_fragment"
20575. 22564
/note="assembly_fragment"
22665. 25349
/note="assembly_fragment"
25450. 28731
/note="assembly_fragment"
28832. 41793
/note="assembly_fragment"
41894. 45562
/note="assembly_fragment"
45663. 49245
/note="assembly_fragment"
49346. 52701
/note="assembly_fragment"
52802. 57587
/note="assembly_fragment"
57688. 60742
/note="assembly_fragment"
60843. 64655

Query Match 3.2%; Score 19; DB 76; Length 156599;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 243 cccgtgttcagccagcc 261
Db 2492 CCGTGTTCAGCCAGCC 2510

RESULT 28
AC015815/c
LOCUS
DEFINITION
Homo sapiens
ACCESSION
AC015815
VERSION
AC015815.4 GI:7341971
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 161549)
Bairren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-21015
Unpublished
2 (bases 1 to 161549)
Bairren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barina,N., Beckerly,R., Boguslavsky,L., Bouckgatter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Gallagan,J., Gardyna,S., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tefaye,S., Tirrell,A., Vassiliev,H., Vo,B., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

TITLE
JOURNAL
Direct Submission
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome

COMMENT

Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 30, 2000 this sequence version replaced gi:6939989.
All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu

Project Information

Center project name: L4073
Center clone name: 21_O_15

Summary Statistics

Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 15133 bases at least Q40
Consensus quality: 155764 bases at least Q30
Consensus quality: 157650 bases at least Q20
Insert size: 157000; agarose-fp
Insert coverage: 4.9 in Q20 bases; agarose-fp
Quality coverage: 4.8 in Q20 bases; sum-of-coverage

NOTE: This is a 'working draft' sequence. It currently consists of 21 contigs. The true order in this sequence record is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 1138: contig of 1138 bp in length
* 1139 1238: gap of 100 bp
* 1239 2713: contig of 1475 bp in length
* 2714 2813: gap of 100 bp
* 2814 4948: contig of 2135 bp in length
* 4949 5048: gap of 100 bp
* 5049 7939: contig of 2891 bp in length
* 7940 8039: gap of 100 bp
* 8040 9902: contig of 1863 bp in length
* 9903 10002: gap of 100 bp
* 10003 12571: contig of 2569 bp in length
* 12572 12671: gap of 100 bp
* 12672 16792: contig of 4121 bp in length
* 16793 16892: gap of 100 bp
* 16893 21320: contig of 4428 bp in length
* 21321 21420: gap of 100 bp
* 21421 28365: contig of 6945 bp in length
* 28366 28465: gap of 100 bp
* 28466 34699: contig of 6234 bp in length
* 34700 34799: gap of 100 bp
* 34800 40251: contig of 5452 bp in length
* 40252 40351: gap of 100 bp
* 40352 47897: contig of 7546 bp in length
* 47898 47997: gap of 100 bp
* 47998 55766: contig of 7769 bp in length
* 55767 55866: gap of 100 bp
* 55867 63968: contig of 8102 bp in length
* 63969 64068: gap of 100 bp
* 64069 74694: contig of 10626 bp in length
* 74695 74794: gap of 100 bp
* 74795 84737: contig of 9943 bp in length
* 84738 84837: gap of 100 bp
* 84838 98320: contig of 13483 bp in length
* 98321 98420: gap of 100 bp
* 98421 110611: contig of 12191 bp in length
* 110612 110711: gap of 100 bp
* 110712 122416: contig of 11707 bp in length
* 122419 122518: gap of 100 bp
* 122519 137004: contig of 14486 bp in length
* 137005 137104: gap of 100 bp

```

FEATURES * 137105 161549: contig of 24445 bp in length.
source Location/Qualifiers

```

1..161549
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-21015"
/clone_11b="RPC1-11 Human Male BAC"
1..1138
misc_feature /note="assembly-fragment"
1239..2713
misc_feature /note="assembly-fragment"
2814..4948
misc_feature /note="assembly-fragment"
5049..7939
misc_feature /note="assembly-fragment"
8040..9902
misc_feature /note="assembly-fragment"
10003..12571
misc_feature /note="assembly-fragment"
12672..16792
misc_feature /note="assembly-fragment"
16893..21320
misc_feature /note="assembly-fragment"
21421..28365
misc_feature /note="assembly-fragment"
28466..34699
misc_feature /note="assembly-fragment"
34800..40251
misc_feature /note="assembly-fragment"
40352..47897
misc_feature /note="assembly-fragment"
47998..55766
misc_feature /note="assembly-fragment"
55867..63968
misc_feature /note="assembly-fragment"
64069..74694
misc_feature /note="assembly-fragment"
74795..84737
misc_feature /note="assembly-fragment"
84838..98320
misc_feature /note="assembly-fragment"
98421..110611
misc_feature /note="assembly-fragment"
110712..122418
misc_feature /note="assembly-fragment"
122519..137004
misc_feature /note="assembly-fragment"
137105..161549
misc_feature /note="assembly-fragment"
BASE COUNT 41638 a 38819 c 39363 g 39720 t 2009 others
ORIGIN

```

Query Match 3.2%: Score 19; DB 64; Length 161549;
Best Local Similarity 100.0%: Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 105 cagggaagaacagcagcca 123
Db 28267 CAGGSAAGACAGCAGCCA 28249

RESULT 29

AC009108 AC009108 170431 bp DNA HTG 02-SEP-2000
LOCUS Homo sapiens chromosome 16 clone RP11-46309, WORKING DRAFT
DEFINITION SEQUENCE, 4 ordered pieces.
ACCESSION AC009108
VERSION AC009108.8 GI:9964740

```

KEYWORDS      HTG: HTGS_PHASE2; HTGS_DRAFT.
SOURCE        human.
ORGANISM      Homo sapiens
              Euxariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE     Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS       1 (bases 1 to 170431)
              DOE Joint Genome Institute.
TITLE         Sequencing of Human Chromosome 16
JOURNAL       Unpublished
REFERENCE     2 (bases 1 to 170431)
              DOE Joint Genome Institute.
AUTHORS       Direct Submission
              Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
COMMENT       Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
              On Sep 2, 2000 this sequence version replaced gi:8575957.
              -----Genome Center
              Center: Joint Genome Institute
              Center Code: JGI
              Web site: http://www.jgi.doe.gov
              -----
Project Information
Center Project Name: 584502
Center Clone name: RPCT-11_46309
-----
Summary Statistics
Consensus quality: 167345 bases at least Q40
Consensus quality: 169423 bases at least Q30
Consensus quality: 169900 bases at least Q20
Estimated insert size: 181440; agarose-fp estimation
Quality coverage: 10.63 in Q20 bases; agarose-fp estimation
Quality coverage: 11.33 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently
consists of 4 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces
is believed to be correct as given, however the sizes
of the gaps between them are based on estimates that have
provided by the submittor.
This sequence will be replaced
by the finished sequence as soon as it is available and
the accession number will be preserved.
1 24820: contig of 24820 bp in length
* 24821 24920: gap of unknown length
* 24921 37862: contig of 12942 bp in length
* 37863 37962: gap of unknown length
* 37963 114671: contig of 76709 bp in length
* 114672 114771: gap of unknown length
* 114772 170431: contig of 55660 bp in length.
Location/Qualifiers
1..170431
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-46309"
/clone_1lb="RPCI human BAC library 11"
BASE COUNT    41117 a 43130 c 42591 g 43290 t          303 others
ORIGIN
Query Match   3.2%; Score 19; DB 61; Length 170431;
Best Local Similarity 100.0%; Prid. No. 26;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 280 gccgcggcgagtcgctg 298
|||||
db 21976 gccgcggcgcgatcgctg 21994
RESULT 30
AC013343/c AC013343 175249 bp DNA HTG 30-MAR-2000
LOCUS      Homo sapiens clone RP11-22B10, WORKING DRAFT SEQUENCE, 13 unordered
DEFINITION pieces.

```

```

AC013343      GI:7341972
AC013343.3   HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS      Homo sapiens
SOURCE        Homo sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE     Birren,B., Linton,L., Nusbaum,C. and Lander,E.
AUTHORS       1 (bases 1 to 175249)
TITLE         Homo sapiens, clone RP11-22B10
JOURNAL       Unpublished
REFERENCE     2 (bases 1 to 175249)
AUTHORS       Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
              Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Bouckgalter,B.,
              Brown,A., Castle,A., Colangelo,M., Collins,S., Collumore,A.,
              Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
              Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
              Galagan,J., Gardyna,S., Grant,G., Hagsos,B., Heatford,A., Horton,L.,
              Howland,J.C., Johnson,R., Jones,C., Kann,L., Karalas,A., Klein,J.J.,
              Lebecky,J., Lien,C., Locke,K., Macdonald,P., Margulis,N.,
              Mcwan,P., McGurk,A., McKernan,K., McDonald,I.P., Meldrim,J.,
              Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
              Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
              Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
              Tesfaye,S., Tirelli,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
              Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
COMMENT       Direct Submission
              Submitted (06-NOV-1999) Whitehead Institute/MIT Center for Genome
              Research, 330 Charles Street, Cambridge, MA 02141, USA
              On Mar 30, 2000 this sequence version replaced gi:6478975.
              All repeats were identified using RepeatMasker:
              Smit, A.F.A. & Green, P. (1996-1997)
              http://ftp.genome.washington.edu/RM/RepeatMasker.html
              -----
              Center: Whitehead Institute/ MIT Center for Genome Research
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@genome.wi.mit.edu
              -----
              Project Information
              Center project name: L4088
              Center Clone name: 22.B_10
              -----
              Summary Statistics
              Sequencing vector: MJ3; M7815; 100% of reads
              Chemistry: Dye-terminator Big Dye; 100% of reads
              Assembly program: Phrap; version 0.960731
              Consensus quality: 161915 bases at least Q40
              Consensus quality: 171919 bases at least Q30
              Consensus quality: 172990 bases at least Q20
              Insert size: 172000; agarose-fp
              Insert size: 174049; sum-of-contigs
              Quality coverage: 6.8 in Q20 bases; agarose-fp
              Quality coverage: 6.7 in Q20 bases; sum-of-contigs
              -----
              NOTE: This is a 'working draft' sequence. It currently
              * consists of 13 contigs. The true order of the pieces
              * is not known and their order in this sequence record is
              * arbitrary. Gaps between the contigs are represented as
              * runs of N, but the exact sizes of the gaps are unknown.
              * This record will be updated with the finished sequence
              * as soon as it is available and the accession number will
              * be preserved.
              -----
              1          1100: contig of 1100 bp in length
              *          1101 1200: gap of 100 bp
              *          1201 2923: contig of 1723 bp in length
              *          2924 3023: gap of 100 bp
              *          3024 3023: contig of 3182 bp in length
              *          6206 6305: gap of 100 bp
              *          6306 8632: contig of 2327 bp in length
              *          8633 8732: gap of 100 bp
              *          8733 16009: contig of 7277 bp in length
              *          16010 16109: gap of 100 bp
              *          16110 26909: contig of 10800 bp in length
              *          26910 27009: gap of 100 bp

```

* 27010 39473: contig of 12464 bp in length
 * 39474 39573: gap of 100 bp
 * 39574 51751: contig of 12178 bp in length
 * 51752 51851: gap of 100 bp
 * 51852 66815: contig of 14964 bp in length
 * 66816 66915: gap of 100 bp
 * 66916 79391: contig of 12476 bp in length
 * 79392 79491: gap of 100 bp
 * 79492 105904: contig of 26413 bp in length
 * 105905 106004: gap of 100 bp
 * 106005 130201: contig of 24197 bp in length
 * 130202 130301: gap of 100 bp
 * 130302 175249: contig of 44948 bp in length.

FEATURES
 source
 Location/Qualifiers
 1. 175249
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="RP11-22B10"
 /clone_11b="RPC1-11 Human Male BAC"
 1. 1100

misc_feature /note="assembly_fragment"
 1201. 2923
 /note="assembly_fragment"
 3024. 6205
 /note="assembly_fragment"
 6306. 8632
 /note="assembly_fragment"
 clone_end:17
 vector_side:right"
 8733. 16009
 /note="assembly_fragment"
 16110. 26909
 /note="assembly_fragment"
 27010. 39473
 /note="assembly_fragment"
 39574. 51751
 /note="assembly_fragment"
 51852. 66815
 /note="assembly_fragment"
 66916. 79391
 /note="assembly_fragment"
 79492. 105904
 /note="assembly_fragment"
 106005. 130201
 /note="assembly_fragment"
 130302. 175249
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left"
 BASE COUNT 55507 a 34489 c 33847 g 50204 t 1202 others
 ORIGIN

Query Match 3.2%; Score 19; DB 63; Length 175249;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 190 agtagcagctatccaata 208
 |||
 Db 68509 AGTAGCAGCTATCCAATA 68491

RESULT 31
 LOCUS AP002501 179844 bp DNA HTG 13-JUN-2000
 DEFINITION Homo sapiens chromosome 18 clone RP11-680K13 map 18q12, WORKING
 DRAFT SEQUENCE, 26 unordered pieces.
 AP002501
 ACCESSION AP002501.1 GI:8547586
 VERSION HTG: HTGS_PHASE1: HTGS_DRAFT.
 KEYWORDS Homo sapiens DNA; clone:RP11-680K13.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 179844)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Homo sapiens 179,844 genomic DNA of 18q12
 Published Only in Database (2000) In press
 2 (bases 1 to 179844)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Direct Submission
 Submitted (09-JUN-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
 Japan (E-mail:hattori@gsc.riken.go.jp,
 URL:http://hgp.gsc.riken.go.jp/,
 Fax:81-42-778-9924)
 ----- Genome Center
 Center: RIKEN Genomic Sciences Center(GSC)
 Center code: RIKEN
 Web site: http://hgp.gsc.riken.go.jp/
 Contact: hattori@gsc.riken.go.jp
 ----- Project Information
 Center project name: HumDrafl18
 Center clone name: RP11-680K13
 ----- Summary Statistics
 Sequencing vector: PCR products; 100% of reads
 Chemistry: Dye-terminator ET-amersham; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 15359 bases at least Q40
 Consensus quality: 16513 bases at least Q30
 Consensus quality: 173786 bases at least Q20
 Insert size: 177344; sum-of-contigs
 Quality coverage: 4.16x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
 26 contigs. The true order of the pieces is not known and their
 order in this sequence record is arbitrary. Gaps between the
 contigs are represented as runs N, but the exact sizes of the gaps
 are unknown. This record will be updated with the finished sequence
 as soon as it is available and the accession number will be
 preserved

1 18994 contig of 18994 bp in length
 19995 39625 contig of 19631 bp in length
 39726 54530 contig of 14805 bp in length
 54631 68239 contig of 13609 bp in length
 68340 78114 contig of 9775 bp in length
 78215 89873 contig of 11659 bp in length
 89974 99848 contig of 9875 bp in length
 99949 109211 contig of 9263 bp in length
 109312 116907 contig of 7596 bp in length
 117008 121661 contig of 4654 bp in length
 121762 128613 contig of 6852 bp in length
 128714 132862 contig of 4149 bp in length
 132963 137321 contig of 4359 bp in length
 137422 142209 contig of 4788 bp in length
 142310 146753 contig of 4444 bp in length
 146854 152208 contig of 5355 bp in length
 152309 158467 contig of 6159 bp in length
 158568 162542 contig of 3908 bp in length
 162643 165710 contig of 3062 bp in length
 165811 168812 contig of 3002 bp in length
 168913 170572 contig of 1660 bp in length
 170673 172454 contig of 1782 bp in length
 172555 175531 contig of 2977 bp in length
 175632 177476 contig of 1845 bp in length
 177577 178671 contig of 1095 bp in length
 178772 179844 contig of 1073 bp in length.

NOTE: This is a 'working draft' sequence. It currently
 consists of 26 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence

* as soon as it is available and the accession number will
* be preserved.

```

1 19894: contig of 19894 bp in length
19895 19994: gap of 100 bp
19995 39625: contig of 19631 bp in length
39626 39725: gap of 100 bp
39726 54530: contig of 14805 bp in length
54531 54630: gap of 100 bp
54631 68239: contig of 13609 bp in length
68240 68339: gap of 100 bp
68340 78114: contig of 9775 bp in length
78115 78214: gap of 100 bp
78215 89873: contig of 11659 bp in length
89874 89973: gap of 100 bp
89974 99848: contig of 9875 bp in length
99849 99948: gap of 100 bp
109212 109311: contig of 9263 bp in length
109312 116907: contig of 7596 bp in length
116908 117007: gap of 100 bp
117008 121661: contig of 4654 bp in length
121662 121761: gap of 100 bp
121762 128613: contig of 6852 bp in length
128614 128713: gap of 100 bp
128714 132862: contig of 4149 bp in length
132863 132962: gap of 100 bp
132963 137321: contig of 4359 bp in length
137322 137421: gap of 100 bp
137422 142209: contig of 4788 bp in length
142210 142309: gap of 100 bp
142310 146753: contig of 4444 bp in length
146754 146853: gap of 100 bp
146854 152208: contig of 5355 bp in length
152209 152308: gap of 100 bp
152309 158467: contig of 6159 bp in length
158468 158567: gap of 100 bp
158568 162542: contig of 3975 bp in length
162543 162642: gap of 100 bp
162643 165710: contig of 3068 bp in length
165711 165810: gap of 100 bp
165811 168812: contig of 3002 bp in length
168813 168912: gap of 100 bp
168913 170572: contig of 1660 bp in length
170573 170672: gap of 100 bp
170673 172454: contig of 1782 bp in length
172455 172554: gap of 100 bp
172555 175531: contig of 2977 bp in length
175532 175631: gap of 100 bp
175632 177476: contig of 1845 bp in length
177477 177576: gap of 100 bp
177577 178671: contig of 1095 bp in length
178672 178771: gap of 100 bp
178772 179844: contig of 1073 bp in length.

```

FEATURES

```

source
1. 179844
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="18"
   /map="18q12"
   /clone="RP11-680K13"
1. 19894
   /note="assembly_fragment"
19895..39625
   /note="assembly_fragment"
39726..54530
   /note="assembly_fragment"
54631..68239
   /note="assembly_fragment"
68340..78114
   /note="assembly_fragment"
78215..89873
   /note="assembly_fragment"
89974..99848

```

```

misc_feature
   /note="assembly_fragment clone_end:SP6 vector_side:left"
99949..109211
   /note="assembly_fragment"
misc_feature
   /note="assembly_fragment"
109312..116907
   /note="assembly_fragment"
117008..121661
   /note="assembly_fragment clone_end:r7 vector_side:right"
121762..128613
   /note="assembly_fragment"
128714..132862
   /note="assembly_fragment"
132963..137321
   /note="assembly_fragment"
137422..142209
   /note="assembly_fragment"
142310..146753
   /note="assembly_fragment"
146854..152208
   /note="assembly_fragment"
152309..158467
   /note="assembly_fragment"
158568..162542
   /note="assembly_fragment"
162643..165710
   /note="assembly_fragment"
165811..168812
   /note="assembly_fragment"
168913..170572
   /note="assembly_fragment"
170673..172454
   /note="assembly_fragment"
172555..175531
   /note="assembly_fragment"
175632..177476
   /note="assembly_fragment"
177577..178671
   /note="assembly_fragment"
178772..179844
   /note="assembly_fragment"
misc_feature
   /note="assembly_fragment"
42717 a 45958 c 47394 g 41275 t 2500 others
ORIGIN

```

```

Query Match      3.28: Score 19; DB 83; Length 179844;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 243 cccgttcagcgcagcc 261
Db 99697 CCTGTTCAGCCGACGC 99679

```

```

RESULT 32
AC068169/c      DNA      15-NOV-2000
LOCUS          AC068169 180717 bp
DEFINITION     Homo sapiens chromosome 17 clone RP11-651B2 map 17, WORKING DRAFT
SEQUENCE       AC068169 15 unordered pieces.
ACCESSION      AC068169.3 GI:11136767
VERSION        HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS       human.
SOURCE         Homo sapiens
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 180717)
AUTHORS        Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE          Homo sapiens chromosome 17, clone RP11-651B2
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 180717)
AUTHORS        Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,

```

TITLE
JOURNAL

COMMENT

Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Gargan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Hearford, A., Horton, L.,
Klein, J., Lacombe, K., Lamazares, R., Jones, C., Kam, L., Karatas, A.,
Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N.,
McCarthy, M., McKean, P., McGurk, A., McKernan, K., McPheters, R.,
Meldrum, J., Menus, L., Minova, C., Miranda, C., Mlenka, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tessier, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, D., Zimmer, A. and Zody, M.

Direct Submission

Submitted (29-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 10, 2000 this sequence version replaced gi:8389515.

All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WtBR

Web site: <http://www.seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L10237

Center clone name: 651.B.2

Summary Statistics

Sequencing vector: M13; M77815; 43% of reads
Sequencing vector: Plasmid; n/a; 57% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 173839 bases at least Q40
Consensus quality: 177110 bases at least Q30
Consensus quality: 178528 bases at least Q20
Insert size: 179000; agarose-fp
Insert size: 179317; sum-of-contigs
Quality coverage: 8.5 in Q20 b.

NOTE: This is a 'working draft' sequence. It currently
consists of 15 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 756: contig of 756 bp in length
757 856: gap of 100 bp
857 2208: contig of 1352 bp in length
2209 2308: gap of 100 bp
2309 3715: contig of 1407 bp in length
3716 3815: gap of 100 bp
3816 5977: contig of 2162 bp in length
5978 6077: gap of 100 bp
6078 57570: contig of 51493 bp in length
57571 57670: gap of 100 bp
57671 60754: contig of 3084 bp in length
60755 60854: gap of 100 bp
60855 65276: contig of 4422 bp in length
65277 65376: gap of 100 bp
65377 71222: contig of 5846 bp in length
71223 71322: gap of 100 bp
71323 81741: contig of 10419 bp in length
81742 81841: gap of 100 bp
81842 91954: contig of 10113 bp in length
91955 92054: gap of 100 bp
92055 104839: contig of 12785 bp in length

104840 104939: gap of 100 bp
104940 125196: contig of 20257 bp in length
125197 125296: gap of 100 bp
125297 149973: contig of 24677 bp in length
149974 150073: gap of 100 bp
150074 178960: contig of 28887 bp in length
178961 179060: gap of 100 bp
179061 180717: contig of 1657 bp in length.

FEATURES

source

1. 180717
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"

misc_feature

/clone="RP11-651B2"
/clone_lib="RP11-651B2 Human Male BAC"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

misc_feature

/note="assembly_fragment"

BASE COUNT

48215 a 43298 c 41795 g 46002 t 1407 others

ORIGIN

Query Match 3.2%; Score 19; DB 73; Length 180717;
Best Local Similarity 100.0%; Pred. No. 26;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 105 caggagagacagcagcagca 123

DB 132375 CAGGAGAGACAGCAGCAGCA 132357

RESULT 33

AC024507

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AC024507 181853 bp DNA HTG 16-MAR-2000
Homo sapiens chromosome 1 clone RP11-153J19 map 1, *** SEQUENCING
IN PROGRESS ***; 55 unordered pieces.

AC024507.2 GI:7249351
HTG: HTGS_PPHASE1.

human.
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 181853)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavsky,L.,
Bouknight,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choeplet,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
DeArliano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M.,
Fenster,J., Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hages,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Landers,T., Larcocque,K., Lehocsky,J., Levine,R.,
Lieu,C., Liu,G., Locke,K., Macdonald,P., Margulis,N., McCarthy,M.,
McEwan,P., McGuirk,A., McKernan,K., McInerney,R., Meldrum,J.,
Meneus,L., Milhova,T., Miranda,C., Mlenka,V., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Olivar,T.M.,
Peterson,K., Pierre,N., Pisan,C., Pollara,V., Raymond,C.,
Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Tirrell,A.,
Travers,M., Triggillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
Wu,X., Wymen,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
Zody,M.

Direct Submission
Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 16, 2000 this sequence version replaced g1:108303.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996.1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7202
Center clone name: 153_J19

NOTE: This is a 'working draft' sequence. It currently
* consists of 55 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1225: contig of 1225 bp in length
1226 1325: gap of 100 bp
1326 2383: contig of 1058 bp in length
2384 2483: gap of 100 bp
2484 3638: contig of 1155 bp in length
3639 3738: gap of 100 bp
3739 4871: contig of 1133 bp in length
4872 4971: gap of 100 bp
4972 6453: contig of 1482 bp in length
6454 6553: gap of 100 bp
6554 7651: contig of 1098 bp in length
7652 7751: gap of 100 bp
7752 8917: contig of 1166 bp in length
8918 9017: gap of 100 bp
9018 10865: contig of 1848 bp in length
10866 10965: gap of 100 bp
10966 12050: contig of 1085 bp in length
12051 12150: gap of 100 bp
12151 13291: contig of 1141 bp in length
13292 13391: gap of 100 bp
13392 14491: contig of 1100 bp in length

14492 14591: gap of 100 bp
14592 15649: contig of 1058 bp in length
15650 15749: gap of 100 bp
15750 17150: contig of 1401 bp in length
17151 17250: gap of 100 bp
17251 18421: contig of 1171 bp in length
18422 18521: gap of 100 bp
18522 20045: contig of 1524 bp in length
20046 20145: gap of 100 bp
20146 21485: contig of 1340 bp in length
21486 21585: gap of 100 bp
21586 23169: contig of 1584 bp in length
23170 23269: gap of 100 bp
23270 24275: contig of 1006 bp in length
24276 24375: gap of 100 bp
24376 25509: contig of 1134 bp in length
25510 25609: gap of 100 bp
25610 26831: contig of 1222 bp in length
26832 26931: gap of 100 bp
26932 27248: contig of 317 bp in length
27249 27348: gap of 100 bp
27349 28595: contig of 1247 bp in length
28596 30616: contig of 1921 bp in length
30617 30716: gap of 100 bp
30717 32613: contig of 1897 bp in length
32614 32713: gap of 100 bp
32714 33958: contig of 1245 bp in length
33959 34058: gap of 100 bp
34059 35732: contig of 1674 bp in length
35733 35832: gap of 100 bp
35833 37422: contig of 1590 bp in length
37423 37522: gap of 100 bp
37523 38306: contig of 784 bp in length
38307 38406: gap of 100 bp
38407 40880: contig of 2474 bp in length
40881 40980: gap of 100 bp
40981 42955: contig of 1975 bp in length
42956 43055: gap of 100 bp
43056 45418: contig of 2363 bp in length
45419 45518: gap of 100 bp
45519 48193: contig of 2675 bp in length
48194 48293: gap of 100 bp
48294 50938: contig of 2645 bp in length
50939 51038: gap of 100 bp
51039 54450: contig of 3412 bp in length
54451 54550: gap of 100 bp
54551 57303: contig of 2753 bp in length
57304 57403: gap of 100 bp
57404 60335: contig of 2932 bp in length
60336 60435: gap of 100 bp
60436 64090: contig of 3655 bp in length
64091 64190: gap of 100 bp
64191 67400: contig of 3210 bp in length
67401 67500: gap of 100 bp
67501 70359: contig of 2859 bp in length
70360 70459: gap of 100 bp
70460 74596: contig of 4137 bp in length
74597 74696: gap of 100 bp
74697 78944: contig of 4248 bp in length
78945 79044: gap of 100 bp
79045 83836: contig of 4792 bp in length
83837 83936: gap of 100 bp
83937 89561: contig of 5625 bp in length
89562 89661: gap of 100 bp
89662 95570: contig of 5909 bp in length
95571 95670: gap of 100 bp
95671 100625: contig of 4955 bp in length
100626 100725: gap of 100 bp
100726 106496: contig of 5771 bp in length
106497 106596: gap of 100 bp
106597 113908: contig of 7312 bp in length
113909 114008: gap of 100 bp

* 114009 119847: contig of 5839 bp in length
* 119848 119947: gap of 100 bp
* 119948 126028: contig of 6081 bp in length
* 126028 126128: gap of 100 bp
* 126129 132814: contig of 6686 bp in length
* 132815 132914: gap of 100 bp
* 132915 139664: contig of 6750 bp in length
* 139665 139764: gap of 100 bp
* 139765 149495: contig of 9731 bp in length
* 149496 149595: gap of 100 bp
* 149596 157948: contig of 8353 bp in length
* 157949 158048: gap of 100 bp
* 158049 167045: contig of 8997 bp in length
* 167046 167145: gap of 100 bp
* 167146 181853: contig of 14708 bp in length.

FEATURES

SOURCE

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1"
/clone="RP11-153J19"
/clone_lib="RPC1-11 Human Male BAC"
1. .1225
misc_feature
/note="assembly_fragment"
1326. .2383
misc_feature
/note="assembly_fragment"
2484. .3638
misc_feature
/note="assembly_fragment"
3739. .4871
misc_feature
/note="assembly_fragment"
4972. .6453
misc_feature
/note="assembly_fragment"
6554. .7651
misc_feature
/note="assembly_fragment"
7752. .8917
misc_feature
/note="assembly_fragment"
9018. .10865
misc_feature
/note="assembly_fragment"
10966. .12050
misc_feature
/note="assembly_fragment"
12151. .13291
misc_feature
/note="assembly_fragment"
13392. .14491
misc_feature
/note="assembly_fragment"
14592. .15649
misc_feature

Query Match

Best Local Similarity 3.2%; Score 19; DB 68; Length 181853;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 484 ttgtgtgcagctaatatgtg 502
|||||

Db 142035 TTGTGTGCAGCTAATATGTG 142053

RESULT 34

LOCUS AC090652 185892 bp DNA HTG 17-APR-2001
DEFINITION Mus musculus clone rp23-116a10, WORKING DRAFT SEQUENCE, 37
unordered pieces.
AC090652
VERSION AC090652.12 GI:13654354
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 185892)
AUTHORS Jiang,X., Song,L. and Roe,B.A.
TITLE Mus musculus BAC Clone rp23-116a10
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 185892)

AUTHORS
TITLE
JOURNAL

COMMENT

Jiang,X., Song,L. and Roe,B.A.
Direct Submission
Submitted (07-MAR-2001) Department Of Chemistry And Biochemistry,
The university Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
Ok 73019, USA
On Apr 17, 2001 this sequence version replaced gi:13592216.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 37 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
2257 2256: contig of 2256 bp in length
2357 2356: gap of unknown length
2357 4894: contig of 2538 bp in length
4895 4994: gap of unknown length
4995 7124: contig of 2130 bp in length
7125 7224: gap of unknown length
7225 9313: contig of 2089 bp in length
9314 9413: gap of unknown length
9414 11675: contig of 2262 bp in length
11676 11775: gap of unknown length
11776 13995: contig of 2220 bp in length
13996 14095: gap of unknown length
14096 16200: contig of 2105 bp in length
16201 16300: gap of unknown length
16301 19350: contig of 3050 bp in length
19351 19451: gap of unknown length
19451 22099: contig of 2649 bp in length
22099 22100: gap of unknown length
22100 22200: gap of unknown length
22200 25395: contig of 3196 bp in length
25396 25496: gap of unknown length
25496 28785: contig of 3290 bp in length
28786 28886: gap of unknown length
28886 31919: contig of 3034 bp in length
31920 32019: gap of unknown length
32020 34426: contig of 2407 bp in length
34427 34526: gap of unknown length
34527 38376: contig of 3850 bp in length
38377 38477: gap of unknown length
38477 41363: contig of 2887 bp in length
41364 41464: gap of unknown length
41464 43857: contig of 2394 bp in length
43857 43958: gap of unknown length
43958 47633: contig of 3676 bp in length
47634 47733: gap of unknown length
47733 51680: contig of 3947 bp in length
51681 51780: gap of unknown length
51781 56459: contig of 4679 bp in length
56460 56559: gap of unknown length
56560 61682: contig of 5123 bp in length
61683 61782: gap of unknown length
61783 68058: contig of 6276 bp in length
68059 68158: gap of unknown length
68159 73468: contig of 5310 bp in length
73469 73568: gap of unknown length
73569 79966: contig of 6398 bp in length
79967 80066: gap of unknown length
80067 85383: contig of 5317 bp in length
85384 85484: gap of unknown length
85484 91630: contig of 6147 bp in length
91631 91730: gap of unknown length
91731 97583: contig of 5853 bp in length
97584 104398: contig of 6715 bp in length
104399 104498: gap of unknown length
104499 111386: contig of 6888 bp in length
111387 111486: gap of unknown length
111487 118341: contig of 6855 bp in length
118342 118442: gap of unknown length
118442 125657: contig of 7216 bp in length

125658 125757: gap of unknown length
* 125758 132358: contig of 6601 bp in length
* 132359 132458: gap of unknown length
* 132459 140201: contig of 7743 bp in length
* 140202 140301: gap of unknown length
* 140302 145956: contig of 5655 bp in length
* 145957 146056: gap of unknown length
* 146057 153347: contig of 7291 bp in length
* 153348 153447: gap of unknown length
* 153448 162532: contig of 9085 bp in length
* 162533 162632: gap of unknown length
* 162633 173314: contig of 10682 bp in length
* 173315 173415: gap of unknown length
* 173415 185892: contig of 12478 bp in length.
Location/Qualifiers
1. 185892
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="tp23-116a10"
BASE COUNT 46626 a 44442 c 44049 g 46981 t 3794 others
ORIGIN

Query Match 3.2%; Score 19; DB 78; Length 185892;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 ggaagacacagcagcagc 126
|||||
DB 64904 GGAAGAACACACACACACG 64922

RESULT 35
LOCUS AC069560
DEFINITION Mus musculus chromosome 2 clone ct7-22f10 strain 129/Sv ES cell
line Cj7, WORKING DRAFT SEQUENCE, 28 unordered pieces.
AC069560 191257 bp DNA HTG 28-MAR-2001
AC069560
VERSION AC069560.17 GI:13470162
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 191257)
Shauli, S., Rahhal, R., Yao, Z., Rithidech, K. and Roe, B.A.
TITLE Mus musculus Chromosome 2 BAC Clone ctb-22f10
JOURNAL Unpublished
AUTHORS Shauli, S., Rahhal, R., Yao, Z., Rithidech, K. and Roe, B.A.
REFERENCE Direct Submission
JOURNAL Submitted (03-JUN-2000) Department of Chemistry and Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT
On Mar 28, 2001, this sequence version replaced gi:11992964.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
* 2130 2129: contig of 2129 bp in length
* 2230 2229: gap of unknown length
* 4411 4410: contig of 2181 bp in length
* 4511 4510: gap of unknown length
* 6828 6827: contig of 2317 bp in length
* 6927 6926: gap of unknown length
* 6928 6927: gap of unknown length
* 9708 9707: gap of unknown length
* 11754 11753: contig of 2046 bp in length
* 11754 11853: gap of unknown length

11854 14087: contig of 2234 bp in length
* 14088 14187: gap of unknown length
* 14188 16271: contig of 2084 bp in length
* 16272 16371: gap of unknown length
* 16372 19053: contig of 2682 bp in length
* 19054 19153: gap of unknown length
* 19154 21327: contig of 2174 bp in length
* 21328 21427: gap of unknown length
* 21428 24279: contig of 2852 bp in length
* 24280 24379: gap of unknown length
* 24380 26574: contig of 2195 bp in length
* 26575 26674: gap of unknown length
* 26675 29819: contig of 3145 bp in length
* 29820 29919: gap of unknown length
* 29920 32658: contig of 2739 bp in length
* 32659 32758: gap of unknown length
* 32759 35340: contig of 2582 bp in length
* 35341 35440: gap of unknown length
* 35441 38430: contig of 2990 bp in length
* 38431 38530: gap of unknown length
* 38431 41853: contig of 3323 bp in length
* 41854 41953: gap of unknown length
* 41954 45151: contig of 3198 bp in length
* 45152 45251: gap of unknown length
* 45252 49105: contig of 3834 bp in length
* 49105 49205: gap of unknown length
* 49206 53586: contig of 4381 bp in length
* 53587 53686: gap of unknown length
* 53687 57668: contig of 3982 bp in length
* 57669 57769: gap of unknown length
* 57769 63349: gap of 5581 bp in length
* 63350 63449: gap of unknown length
* 63450 67562: contig of 4113 bp in length
* 67563 74350: gap of unknown length
* 74351 74450: contig of 6688 bp in length
* 74451 78798: gap of unknown length
* 78799 78898: contig of 4348 bp in length
* 78899 78899: gap of unknown length
* 78899 84605: contig of 5707 bp in length
* 84606 84705: gap of unknown length
* 84706 99565: contig of 14860 bp in length
* 99566 99565: gap of unknown length
* 99566 136984: gap of unknown length
* 136985 137094: contig of 37329 bp in length
* 137095 191257: contig of 54183 bp in length.
Location/Qualifiers
1. 191257
/organism="Mus musculus"
/strain="129/Sv ES cell line Cj7"
/db_xref="taxon:10090"
/chromosome="2"
/clone="ct7-22f10"
/clone_lib="Caltech CTB-BAC library"
BASE COUNT 54814 a 40040 c 40978 g 52695 t 2730 others
ORIGIN

Query Match 3.2%; Score 19; DB 74; Length 191257;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaaacacagcagcagc 555
|||||
DB 130319 GAAACCAACAGAGAGAGAA 130337

RESULT 36
LOCUS AL512653
DEFINITION Homo sapiens chromosome 1 clone RP11-523k4, *** SEQUENCING IN
PROGRESS ***
ACCESSION AL512653
VERSION AL512653.4 GI:13235023

KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 194575)
 AUTHORS McIay, K.
 TITLE Direct Submission
 JOURNAL Submitted (28-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 Requests: clonerequests@sanger.ac.uk
 On Mar 5, 2001 this sequence version replaced gi:13169575.
 COMMENT
 Sanger Centre
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquerry@sanger.ac.uk
 Project Information
 Center project name: ba523k4
 Summary Statistics
 Sequencing program: XGAP4; version 4.5
 Chemistry: Dye-terminator Big Dye; 98% of reads
 Dye-terminator Big Dye; 98% of reads
 Consensus quality: 171678 bases at least Q40
 Consensus quality: 180206 bases at least Q30
 Consensus quality: 185635 bases at least Q20
 Insert size: 190375; sum-of-contigs
 Insert size: 152230; 3.3% error; agarose-fp
 Quality coverage: 3.14x in Q20 bases; sum-of-contigs Quality
 coverage: 4.72x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 43 contigs. The true order of the pieces is
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 4458: contig of 4458 bp in length
 * 4459 4558: gap of 100 bp
 * 4559 6602: contig of 2044 bp in length
 * 6603 6702: gap of 100 bp
 * 6703 9078: contig of 2376 bp in length
 * 9079 9178: gap of 100 bp
 * 9179 12224: contig of 3046 bp in length
 * 12225 12324: gap of 100 bp
 * 12325 15579: contig of 3255 bp in length
 * 15580 15679: gap of 100 bp
 * 15680 22986: contig of 7307 bp in length
 * 22987 23086: gap of 100 bp
 * 23087 25938: contig of 2852 bp in length
 * 25939 26038: gap of 100 bp
 * 26039 38837: contig of 12799 bp in length
 * 38838 38937: gap of 100 bp
 * 38938 45337: contig of 6400 bp in length
 * 45338 45437: gap of 100 bp
 * 45438 49342: contig of 3905 bp in length
 * 49343 49442: gap of 100 bp
 * 49443 61905: contig of 12463 bp in length
 * 61906 62005: gap of 100 bp
 * 62006 64568: contig of 2563 bp in length
 * 64569 64668: gap of 100 bp
 * 64669 67376: contig of 2708 bp in length
 * 67377 67476: gap of 100 bp
 * 67477 69819: contig of 2343 bp in length
 * 69820 69919: gap of 100 bp
 * 69920 76132: contig of 6213 bp in length
 * 76133 76232: gap of 100 bp
 * 76233 78236: contig of 2004 bp in length
 * 78237 78336: gap of 100 bp
 * 78337 86544: contig of 8208 bp in length

* 86545 86644: gap of 100 bp
 * 86645 89369: contig of 2725 bp in length
 * 89370 89469: gap of 100 bp
 * 89470 93493: contig of 4024 bp in length
 * 93494 93593: gap of 100 bp
 * 93594 96676: contig of 3083 bp in length
 * 96677 96776: gap of 100 bp
 * 96777 99971: contig of 3195 bp in length
 * 99972 100071: gap of 100 bp
 * 100072 105701: contig of 5630 bp in length
 * 105702 105801: gap of 100 bp
 * 105802 109322: contig of 3521 bp in length
 * 109323 109422: gap of 100 bp
 * 109423 112011: contig of 2589 bp in length
 * 112012 112111: gap of 100 bp
 * 112112 114284: contig of 2173 bp in length
 * 114285 114384: gap of 100 bp
 * 114385 122566: contig of 8182 bp in length
 * 122567 122666: gap of 100 bp
 * 122667 129674: contig of 7008 bp in length
 * 129675 129774: gap of 100 bp
 * 129775 132589: contig of 2815 bp in length
 * 132590 132689: gap of 100 bp
 * 132690 134777: contig of 2088 bp in length
 * 134778 134877: gap of 100 bp
 * 134878 141445: contig of 6568 bp in length
 * 141446 141545: gap of 100 bp
 * 141546 149107: contig of 7562 bp in length
 * 149108 149207: gap of 100 bp
 * 149208 152744: contig of 3537 bp in length
 * 152745 152844: gap of 100 bp
 * 152845 158786: contig of 5942 bp in length
 * 158787 158886: gap of 100 bp
 * 158887 161231: contig of 2345 bp in length
 * 161232 161331: gap of 100 bp
 * 161332 163923: contig of 2592 bp in length
 * 163924 164023: gap of 100 bp
 * 164024 167292: contig of 3269 bp in length
 * 167293 167392: gap of 100 bp
 * 167393 169608: contig of 2216 bp in length
 * 169609 169708: gap of 100 bp
 * 169709 175744: contig of 6036 bp in length
 * 175745 175844: gap of 100 bp
 * 175845 180590: contig of 4746 bp in length
 * 180591 180690: gap of 100 bp
 * 180691 183185: contig of 2495 bp in length
 * 183186 183285: gap of 100 bp
 * 183286 185116: contig of 2231 bp in length
 * 185117 185616: gap of 100 bp
 * 185617 192097: contig of 6481 bp in length
 * 192098 192197: gap of 100 bp
 * 192198 194575: contig of 2378 bp in length.
 * Location/Qualifiers
 * 1..194575
 * /organism="Homo sapiens"
 * /db_xref="taxon:9606"
 * /chromosome="1"
 * /clone="RP11-523K4"
 * /clone_11b="RP11-11.2"
 * 1..4458
 * /note="assembly-fragment:00856
 * fragment_chain:1"
 * 4459..6602
 * /note="assembly-fragment:00764
 * fragment_chain:1"
 * 6703..9078
 * /note="assembly-fragment:01170
 * fragment_chain:1"
 * 9179..12224
 * /note="assembly-fragment:00852
 * fragment_chain:1"
 * 12325..15579
 * /note="assembly-fragment:01087

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

FEATURES
source

| | | | | | | | | | |
|-----------------------|---|-------------------|--------------------------------|----------------|-------------|--------|----|------|----|
| misc_feature | fragment_chain:2" | 15680..22986 | /note="assembly_fragment:00246 | | | | | | |
| misc_feature | fragment_chain:2" | 23087..25938 | /note="assembly_fragment:01976 | | | | | | |
| misc_feature | fragment_chain:2" | 26039..38837 | /note="assembly_fragment:01200 | | | | | | |
| misc_feature | fragment_chain:2" | 38938..45337 | /note="assembly_fragment:00679 | | | | | | |
| misc_feature | fragment_chain:3" | 45438..49342 | /note="assembly_fragment:01991 | | | | | | |
| misc_feature | fragment_chain:3" | 49443..61905 | /note="assembly_fragment:00239 | | | | | | |
| misc_feature | fragment_chain:3" | 62006..64568 | /note="assembly_fragment:00334 | | | | | | |
| misc_feature | fragment_chain:4" | 64659..67376 | /note="assembly_fragment:00402 | | | | | | |
| misc_feature | fragment_chain:4" | 67477..69819 | /note="assembly_fragment:00376 | | | | | | |
| misc_feature | fragment_chain:5" | 69920..76132 | /note="assembly_fragment:01449 | | | | | | |
| misc_feature | fragment_chain:5" | 76233..78236 | /note="assembly_fragment:00418 | | | | | | |
| misc_feature | fragment_chain:6" | 78337..86544 | /note="assembly_fragment:00432 | | | | | | |
| misc_feature | fragment_chain:6" | 86645..89369 | /note="assembly_fragment:00503 | | | | | | |
| misc_feature | fragment_chain:7" | 89470..93493 | /note="assembly_fragment:00747 | | | | | | |
| misc_feature | fragment_chain:7" | 93594..96676 | /note="assembly_fragment:00956 | | | | | | |
| misc_feature | fragment_chain:8" | 96777..99971 | /note="assembly_fragment:00401 | | | | | | |
| misc_feature | fragment_chain:8" | | | | | | | | |
| Query Match | 3.2%; | Score 19; | DB 81; | Length 194575; | | | | | |
| Best Local Similarity | 100.0%; | Pred. No. 26; | | | | | | | |
| Matches | 19; | Conservative | 0; | Mismatches | 0; | Indels | 0; | Gaps | 0; |
| QY | 547 | 99aagaagaagcctcg | 565 | | | | | | |
| | | Db 187810 | GGAGAGAAAGCCTCG | 187828 | | | | | |
| RESULT | 37 | | | | | | | | |
| AC011359 | 198453 | bp | DNA | HTG | 14-FEB-2001 | | | | |
| LOCUS | Homo sapiens chromosome 5 | clone CTC-370H24, | WORKING DRAFT | SEQUENCE, | | | | | |
| DEFINITION | 6 ordered pieces. | | | | | | | | |
| ACCESSION | AC011359 | | | | | | | | |
| VERSION | AC011359.4 | GI:12830128 | | | | | | | |
| KEYWORDS | HTG; HTGS_PHASE2; HTGS_DRAFT. | | | | | | | | |
| SOURCE | human. | | | | | | | | |
| ORGANISM | Homo sapiens | | | | | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | | | | | |
| | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | | | | | |
| REFERENCE | 1 (bases 1 to 198453) | | | | | | | | |
| AUTHORS | DOE Joint Genome Institute. | | | | | | | | |

```

TITLE                               Unpublished
JOURNAL                             Sequencing of Human Chromosome 5
REFERENCE                           2 (bases 1 to 198453)
AUTHORS                            DOE Joint Genome Institute.
TITLE                               Direct Submission
JOURNAL                             Submitted (06-OCT-1999) Production Sequencing Facility, DOE Joint
                                   Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT                             On Feb 14, 2001 this sequence version replaced gi:7710570.
                                   -----Genome Center
                                   Center: Joint Genome Institute
                                   Center Code: JGI
                                   Web site: http://www.jgi.doe.gov
                                   -----
Project Information
Center Project Name: 327450, H224
Center clone name: CIT-HSPC_370H24
                                   -----
Summary Statistics
Consensus quality: 192637 bases at least Q40
Consensus quality: 195977 bases at least Q30
Consensus quality: 196983 bases at least Q20
Estimated insert size: 196000; pulse-field gel estimation
Estimated insert size: 197953; sum-of-contigs estimation
Quality coverage: 8.4 in Q20 bases; pulse field gel estimation
Quality coverage: 8.32 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
      *   1    80992: contig of 80992 bp in length
      *   *   80993    81092: gap of unknown length
      *   *   81093    98177: contig of 17085 bp in length
      *   *   98178    98277: gap of unknown length
      *   *   98278    113101: contig of 14824 bp in length
      *   *   113102    113201: gap of unknown length
      *   *   113202    132185: contig of 18984 bp in length
      *   *   132186    132285: gap of unknown length
      *   *   132286    157258: contig of 24973 bp in length
      *   *   157259    157358: gap of unknown length
      *   *   157359    198453: contig of 41095 bp in length.
FEATURES
source
Location/Qualifiers
      1..198453
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="5"
      /clone="CIT-370H24"
BASE COUNT      54938 a 39611 c 41079 g 62319 t      506 Others
ORIGIN
Query Match          3.2% Score 19; DB 62; Length 198453;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      484      ttggtcgcagctaatactg 502
              |||
              |||
              |||
Db 134128 TTGGTGCAGCAATATGTC 134146
RESULT 38
AL445211 AL445211 200860 bp DNA HTG 06-MAR-2001
LOCUS Homo sapiens chromosome 1 clone RP5-882B12, *** SEQUENCING IN
DEFINITION PROGRESS ***, 38 unordered pieces.
ACCESSION AL445211 GI:13273841
VERSION AL445211.5
KEYWORDS HTG; HTGS_PHASE1.
```

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 200860)
 AUTHORS McIlroy, K.
 TITLE Direct Submission
 JOURNAL Submitted (03-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone
 requests: clonerequests@sanger.ac.uk
 On Mar 11, 2001 this sequence version replaced gi:13234984.
 ----- Genome Center
 Center: Sanger Centre
 Web site: http://www.sanger.ac.uk
 Contact: humquerry@sanger.ac.uk
 ----- Project Information
 Center project name: d082812
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Sequencing vector: plasmid; 108752; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Consensus quality: 17636 bases at least Q40
 Consensus quality: 186827 bases at least Q30
 Consensus quality: 192595 bases at least Q20
 Insert size: 197160; sum-of-contigs
 Insert size: 121820; 3.4% error; agarose-fp
 Quality coverage: 2.89x in Q20 bases; sum-of-contigs Quality
 coverage: 5.60x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 38 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 4569: contig of 4569 bp in length
 * 4570 4669: gap of 100 bp
 * 9273 9372: contig of 4603 bp in length
 * 9373 15080: contig of 5708 bp in length
 * 15081 15180: gap of 100 bp
 * 15181 18858: contig of 3678 bp in length
 * 18859 18958: gap of 100 bp
 * 18959 23944: contig of 4986 bp in length
 * 23945 24044: gap of 100 bp
 * 24045 27176: contig of 3132 bp in length
 * 27177 27276: gap of 100 bp
 * 27277 31044: contig of 3768 bp in length
 * 31045 31144: gap of 100 bp
 * 31145 41645: contig of 10501 bp in length
 * 41646 41745: gap of 100 bp
 * 41746 44420: contig of 2675 bp in length
 * 44421 44520: gap of 100 bp
 * 44521 51070: contig of 6550 bp in length
 * 51071 51170: gap of 100 bp
 * 51171 56433: contig of 5269 bp in length
 * 56440 56539: gap of 100 bp
 * 56540 62931: contig of 6392 bp in length
 * 62932 63031: gap of 100 bp
 * 63032 66281: contig of 3250 bp in length
 * 66282 66381: gap of 100 bp
 * 66382 71591: contig of 5210 bp in length
 * 71592 71691: gap of 100 bp
 * 71692 74032: contig of 2341 bp in length
 * 74033 74132: gap of 100 bp
 * 74133 78618: contig of 4486 bp in length
 * 78619 78718: gap of 100 bp
 * 78719 85194: contig of 6476 bp in length
 * 85195 85294: gap of 100 bp
 * 85295 90909: contig of 5615 bp in length

FEATURES
 source
 1..200860
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP5-882B12"
 /clone_lib="RPCT-5"
 1..4569
 /note="assembly_fragment:00173
 fragment_chain:1"
 4670..9272
 /note="assembly_fragment:00415
 fragment_chain:1"
 9373..15080
 /note="assembly_fragment:00424
 fragment_chain:1"
 15181..18858
 /note="assembly_fragment:01264
 fragment_chain:1"
 18959..23944
 /note="assembly_fragment:00053
 fragment_chain:2"
 24045..27176
 /note="assembly_fragment:00723
 fragment_chain:2"
 27277..31044
 /note="assembly_fragment:00266
 fragment_chain:2"
 31145..41645
 /note="assembly_fragment:00802
 fragment_chain:3"
 41746..44420
 /note="assembly_fragment:01634
 fragment_chain:3"

```

misc_feature      fragment_chain:3"
                  44521..51070
                  /note="assembly_fragment:00827
fragment_chain:4"
misc_feature      51171..56439
                  /note="assembly_fragment:01185
fragment_chain:4"
misc_feature      56540..62931
                  /note="assembly_fragment:01751
fragment_chain:5"
misc_feature      63032..66281
                  /note="assembly_fragment:01248
fragment_chain:5"
misc_feature      66382..71591
                  /note="assembly_fragment:01797
fragment_chain:6"
misc_feature      71692..74032
                  /note="assembly_fragment:01025
fragment_chain:6"
misc_feature      74133..78618
                  /note="assembly_fragment:01903
fragment_chain:7"
misc_feature      78719..85194
                  /note="assembly_fragment:00657
fragment_chain:7"
misc_feature      85295..90909
                  /note="assembly_fragment:00240"
misc_feature      91010..94589
                  /note="assembly_fragment:00258"
misc_feature      94690..101005
                  /note="assembly_fragment:00263"
misc_feature      101106..113385
                  /note="assembly_fragment:00313"
misc_feature      113486..116348
                  /note="assembly_fragment:00541"
misc_feature      116449..118982
                  /note="assembly_fragment:00582"
misc_feature      119083..121185
                  /note="assembly_fragment:00590"
misc_feature      121286..129849
                  /note="assembly_fragment:00612"
misc_feature      129950..133006
                  /note="assembly_fragment:00655"
misc_feature      133107..141678
                  /note="assembly_fragment:00725"
misc_feature      141779..150289
                  /note="assembly_fragment:00770"
misc_feature      150390..156647

```

Query Match 3.28; Score 19; DB 81; Length 200860;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 547 ggagagaagaagcctccg 565
 ||||||||||||||||
 Db 113569 GGAGAGAGAAAGCCTCCG 113587

RESULT 39
 AC004615 AC004615 235141 bp DNA PRI 03-FEB-2000
 LOCUS Homo sapiens BAC clone GSI-405L21 from 5p15.2, complete sequence.
 DEFINITION
 AC004615 AC004615.1 GI:3080660
 VERSION
 AC004615.1 GI:3080660
 KEYWORDS
 HTG.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 235141)
 AUTHORS Kallunki,J. and Harmon,G.
 TITLE The sequence of Homo sapiens BAC clone GSI-405L21
 JOURNAL Unpublished

REFERENCE
 2 (bases 1 to 235141)
 AUTHORS
 WATERSTON,R.
 TITLE Direct Submission
 JOURNAL Submitted (24-APR-1998) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 3 (bases 1 to 235141)
 REFERENCE
 3 (bases 1 to 235141)
 AUTHORS
 WATERSTON,R.
 TITLE Direct Submission
 JOURNAL Submitted (03-FEB-2000) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: saplens@watson.wustl.edu
 ----- Summary Statistics

 Center project name: H_GS405L21

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded or sequenced with an alternate
 chemistry; an attempt was made to resolve all sequencing problems,
 such as compressions and repeats; all regions were covered by
 sequence from more than one subclone; and the assembly was
 confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. Michael
 Lovett, Departments of Otorhinolaryngology, Molecular Biology and
 Oncology, University of Texas Southwestern Medical Center, Dallas
 TX. For additional information about the map position of this
 sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
 This clone is from the first BAC library from Genome Systems, Inc.
 (http://www.genomesystems.com).
 Cell line: Lymphoblastoid
 Haplotypes: two
 VECTOR: pBelOBAC
 Selection: chloramphenicol
 NEIGHBORING SEQUENCE INFORMATION:
 The actual start of this clone is at base position 1 of GSI-405L21;
 actual end is at 235141 of GSI-405L21.

FEATURES
 This clone contains STS HSA0552D9 (NID:q1233025).
 Location/Qualifiers

1..235141
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /map="5p15.2"
 /clone="GSI-405L21"
 /clone_1ib="GSBAC1"
 /rpt_family="Alu"
 181..480
 /rpt_family="Alu"
 3000..3303
 /rpt_family="Alu"
 4356..4559
 /rpt_family="Alu"
 4576..4744
 /rpt_family="(TAGA)n"
 6769..7074
 /rpt_family="MER1_type"
 7096..7157
 /rpt_family="Alu"
 7286..7384
 /rpt_family="(CA)n"
 7491..7742
 /rpt_family="MALR"

gene

CDS

/rpt_family="MaLR"
9154..>234415
/gene="SEMA6"
join(9154..9277,51204..51303,70597..70642,151152..151214,
162076..162174,164155..164368,186819..187104,
191786..191921,198574..198778,234207..>234415)
/gene="SEMA7"
/note="match to U52840 (PID:g27722584);but note
polymorphism at position 56; H_GS045121.1"
/codon_start=1
/product="semaphorin F"
/protein_id="AAC14668.1"
/db_xref="GI:3080661"
/translation="MKGTCVIAWLFSSGLMRLAHPAQAQTTOCORTPEYISKEIG
PMLREFPKNAVDPSLTDEPDQKELVVGARVLYERLOLESLIAENHEDCAPTK
ACYSRGSKEBCQNTIRVLVGGDLTPCGRAFPVPTVNSSLTUEPIHQISMAR
CPYSPHNSTALTAAGBELYAATANDPFGRDPAIYRSGLPLPLETAOYNRWLERN
NCSRGDEIFYFEFFRENAAVEHDCKTVFSRAARVCNDIGRFLTEDTWTFPMKAR
FYQQNSNSAMLPNPNDHFCCGTDDGLVNLERNLDQAOKFILHNEVPQPTYV
PFEMDNSSSHVAVDVOGREALVHIILAYNDVCTIKKVRYPLMNQTSLSLLREIEL
PERRRREPRLSQILHSQSYLEFVGLREHVNIPLKRQCFYTRRR"
10974..11242
/rpt_family="L1"
11271..11381
/rpt_family="L1"
12176..12392
/rpt_family="L1"
13345..13391
/rpt_family="MER81"
13453..13898
/rpt_family="MaLR"
19001..19202
/rpt_family="L2"
19350..19841
/rpt_family="L1"
19893..20700
/rpt_family="L1"
21002..21401
/rpt_family="MaLR"
21797..21994
/rpt_family="MER1_type"
22413..22710
/rpt_family="Alu"
27381..27411
/rpt_family="Mariner"
27785..28082
/rpt_family="Alu"
29733..29863
/rpt_family="MIR"
30444..30532
/rpt_family="L2"
33326..33653
/rpt_family="L2"
33961..34048
/rpt_family="MER1_type"
34057..34316
/rpt_family="L2"
34390..34520
/rpt_family="L2"
36072..36286
/rpt_family="MER1_type"
36464..36604
/rpt_family="Alu"
36622..36732
/rpt_family="(TAGA)n"
38061..38641
/rpt_family="MER21_g"
38939..39082
/rpt_family="MER1_type"
39096..39399
/rpt_family="Alu"
39415..39483

| | |
|---------------|--|
| repeat_region | /rpt_family="MER1_type" 39557, .39663 /rpt_family="L2" |
| repeat_region | 42462, .42767 /rpt_family="Retroviral" |
| repeat_region | 42862, .43091 /rpt_family="Retroviral" |
| repeat_region | 46106, .46225 /rpt_family="MIR" |
| repeat_region | 47785, .47839 /rpt_family="(CA)n" |
| repeat_region | 51495, .51731 /rpt_family="MIR" |
| repeat_region | 53061, .53183 /rpt_family="L2" |
| repeat_region | 54138, .54238 /rpt_family="MER53" |
| repeat_region | 54275, .54314 /rpt_family="L2" |
| repeat_region | 54348, .54495 /rpt_family="MIR" |
| repeat_region | 58611, .58868 /rpt_family="Alu" |
| repeat_region | 59061, .59261 /rpt_family="MER2_type" |
| repeat_region | 59469, .59532 /rpt_family="L2" |
| repeat_region | 59571, .59768 /rpt_family="MER1_type" |
| repeat_region | 60224, .60521 /rpt_family="Alu" |
| repeat_region | 62226, .62479 /rpt_family="Alu" |
| repeat_region | 62480, .62757 /rpt_family="Alu" |
| repeat_region | 63954, .64048 /rpt_family="MIR" |
| repeat_region | 64650, .64950 /rpt_family="Alu" |
| repeat_region | 65069, .65357 /rpt_family="Alu" |

| | | | | | | | |
|------------------------|--------|----------------------|-------|------------|----|--------|--------|
| Query Match Similarity | 3.2% | Score | 19 | DB | 85 | Length | 255141 |
| Best Local Similarity | 100.0% | Pred. No. | 25 | | | | |
| Matches | 19 | Conservative | 0 | Mismatches | 0 | Indels | 0 |
| OY | 160 | gggtgtgtgtctctcgggca | 178 | | | | |
| | | | | | | | |
| db | 96608 | gggtgtgtgtctctcgggca | 96626 | | | | |

```

RESULT 40
AL354813
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
-----
AL354813 296050 bp DNA HTG 13-APR-2001
Homo sapiens chromosome 20 clone CTD-2653D5, *** SEQUENCING IN
PROGRESS ***, 14 unordered pieces.
AL354813
AL354813.25 GI:13624971
HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 296050)
Wallis,J.
Direct Submission
Submitted (12-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: hinxton@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Apr 14, 2001 this sequence version replaced gi:13568011.
-----
Genome Center
Center: Sanger Centre
Center code: SC

```


AUTHORS Paranhos-Baccala,G., Lesenechal,M. and Jolivet,M.
 TITLE NOVEL TRYPAOSOMA CRUZI ANTIGEN, AND GENE CODING THEREFOR; THEIR
 APPLICATION TO THE DETECTION OF CHAGAS' DISEASE
 JOURNAL Patent: WO 9605312-A 8 22-FEB-1996;

COMMENT BIO MERIEUX (FR)
 Other publication CA 2173957 960222
 Other publication AU 3169195 960307
 Other publication FR 2723589 960216.

FEATURES
 source 1..18
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 5 a 4 c 7 g 2 t
 ORIGIN

Query Match 3.0%; Score 18; DB 9; Length 18;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 actctgcgcgtcgtcga 228
 Db 18 ACTCTGCCGCTGCTGCA 1

RESULT 43
 LOCUS AR047923 18 bp DNA PAT 29-SEP-1999
 DEFINITION Sequence 5 from patent US 5820864.
 ACCESSION AR047923
 VERSION AR047923.1 GI:5970266
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 18)
 Paranhos-Baccala,G., Lesenechal,M. and Jolivet,M.
 Trypanosoma cruzi antigen, gene encoding therefor and methods of
 detecting and treating Chagas disease
 JOURNAL Patent: US 5820864-A 5 13-OCT-1998;
 FEATURES Location/Qualifiers
 source 1..18

BASE COUNT 2 a 6 c 8 g 2 t
 ORIGIN

Query Match 3.0%; Score 18; DB 9; Length 18;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 tcgggcactgaacggcg 189
 Db 1 TCGGCGACTGACGCGCG 18

RESULT 44
 LOCUS AR047926 18 bp DNA PAT 29-SEP-1999
 DEFINITION Sequence 8 from patent US 5820864.
 ACCESSION AR047926
 VERSION AR047926.1 GI:5970269
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 18)
 Paranhos-Baccala,G., Lesenechal,M. and Jolivet,M.
 Trypanosoma cruzi antigen, gene encoding therefor and methods of
 detecting and treating Chagas disease
 JOURNAL Patent: US 5820864-A 8 13-OCT-1998;
 FEATURES Location/Qualifiers
 source 1..18

BASE COUNT 5 a 4 c 7 g 2 t
 ORIGIN

Query Match 3.0%; Score 18; DB 9; Length 18;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 actctgcgcgtcgtcga 228
 Db 18 ACTCTGCCGCTGCTGCA 1

RESULT 45
 LOCUS AU049339 339 bp DNA STS 20-JAN-2000
 DEFINITION Rattus norvegicus, OTSUKA clone, 862e02, microsatellite sequence,
 sequence tagged site.
 ACCESSION AU049339
 VERSION AU049339.1 GI:6722510
 KEYWORDS STS.
 SOURCE Rattus norvegicus (strain:Brown Norway) liver hepatocyte DNA,
 clone:862e02.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (sites)
 Watanabe,T.K., Hishigaki,H., Okuno,S., Mizoguchi,A., Oga,K.,
 Tsuji,A., Ono,T., Yamasaki,Y., Kanemoto,N., Takahashi,E., Irie,Y.,
 Nakamura,Y., Takagi,Y. and Tanigami,A.
 The large-scale mapping of rat microsatellite markers
 JOURNAL Unpublished (1998)
 AUTHORS 2 (bases 1 to 339)
 TITLE Watanabe,T.K.
 JOURNAL Direct Submission
 Submitted (11-DEC-1998) to the DDBJ/EMBL/GenBank databases. Takeshi
 K Watanabe, Otsuka GEN Research Institute, Otsuka Pharmaceutical
 Co., Ltd; 463-10, Kagasuno, Kawauchi-cho, Tokushima, Tokushima
 771-0192, Japan (E-mail:watanabe@otsuka.gr.jp, Tel:+81-886-65-2888,
 Fax:+81-886-37-1035)

FEATURES
 source 1..339
 /organism="Rattus norvegicus"
 /strain="Brown Norway"
 /db_xref="taxon:10116"
 /cell_type="hepatocyte"
 /clone="862e02"
 /tissue_type="liver"
 /note="862e02F-5'-TCGGTGCACCTGAGTTCAAT-3',
 862e02R-5'-GACATGACACCCACATCATACCA-3'."

BASE COUNT 61 a 73 c 76 g 128 t
 ORIGIN

Query Match 3.0%; Score 18; DB 53; Length 339;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaaccacagagagaaga 554
 Db 23 GAACCAACAGAGAGAGA 6

RESULT 46
 LOCUS AX086831 1589 bp DNA PAT 09-MAR-2001
 DEFINITION Sequence 783 from Patent WO0112659.
 ACCESSION AX086831
 VERSION AX086831.1 GI:13276044
 KEYWORDS human.
 SOURCE

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1589)
 AUTHORS Wiemann, S.
 TITLE Human dna sequences
 JOURNAL Patent: WO 0112659-A 783 22-FEB-2001;
 German Human Genome Project (DE)
 FEATURES
 Source Location/Qualifiers
 1. 1589
 /db_xref="taxon:9606"
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 442 a 374 c 353 g 420 t
 ORIGIN

Query Match 3.0%; Score 18; DB 10; Length 1589;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 actctgcgcgtctgca 228
 ||||||||||||||||
 Db 333 ACTTCTGCCGCTGCTCA 316

RESULT 47
 HSM801826/c
 LOCUS HSM801826 1589 bp mRNA PRI 10-MAR-2001
 DEFINITION Homo sapiens mRNA: cDNA DKFZp434N2435 (from clone DKFZp434N2435);
 complete cds.
 ACCESSION AL136858
 VERSION AL136858.1 GI:12053220
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1589)
 AUTHORS Wiemann, S., Weill, B., Wellenreuther, R., Gassenhuber, J., Glassl, S.,
 Ansorge, W., Boecker, M., Bloeker, H., Bauersachs, S., Blum, H.,
 Lauber, D., Duesterhoeft, A., Beyer, A., Koehrer, K., Strack, N.,
 Mewes, H.W., Ottenwelder, B., Obermaler, B., Tampe, J., Heubner, D.,
 Wambutt, R., Korn, B., Klein, M. and Poustka, A.
 TITLE Toward a Catalog of Human Genes and Proteins: Sequencing and
 Analysis of 500 Novel Complete Protein Coding Human cDNAs
 JOURNAL Genome Res. 11 (3), 422-435 (2001)
 PUBMED 11230166
 REFERENCE 2 (bases 1 to 1589)
 AUTHORS Poustka, A., Klein, M., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
 TITLE Direct Submission
 JOURNAL Submitted (18-JAN-2000) MIPS, Am Klopferstr. 18a, D-82152
 Martinsried, GERMANY
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
 Heidelberg/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 This clone is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at http://www.mips.blochem.mpg.de/proj/cDNA/
 Location/Qualifiers
 1. 1589
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZp434N2435"
 /tissue.type="testis"
 /clone_lib="434 (synonym: htess3). Vector psport1; host
 DH10B; sites NotI + SalI"
 /dev_stage="adult"
 78. 1175
 /gene="DKFZp434N2435"

CDS 78. 1175
 /gene="DKFZp434N2435"
 /note="unknown protein"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CA66792.1"
 /db_xref="GI:12053221"
 /translation="MNVLYPLAVPKGRRLCEVCEAPAEAVCAACTVITYCGVHOKA
 DMSIHEKICQLILPLRTSMPEYNSEERQHQLOQROKYLIEFCYTIARQYLIEG
 KHEDAVPAALQSLRFKLVGLSSVELVAYPLPAASIGLRIYVAEYLFQAOQMTV
 LKSTDCSNATHSLIHRNIGILYIARKNVEARVHLANDYFSCACGTDIRSGCYF
 HLANTFYDKLADLTLYTKVSEIYHAILNNHYQVLSAHIQOQNLKLEPNDGL
 DEQDEARITLITLIRNDSIDKAPQKTFVLYLKLTVMLYTLMMNSRAQETGMAL
 SLAEQQLVHEQSTIQELISTEDHPIT"
 1561. 1566
 polyA_signal
 polyA_site
 BASE COUNT 442 a 374 c 353 g 420 t
 ORIGIN

Query Match 3.0%; Score 18; DB 93; Length 1589;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 actctgcgcgtctgca 228
 ||||||||||||||||
 Db 333 ACTTCTGCCGCTGCTCA 316

RESULT 48
 AR051480
 LOCUS AR051480 2520 bp DNA PAT 29-SEP-1999
 DEFINITION Sequence 50 from patent US 5830670.
 ACCESSION AR051480
 VERSION AR051480.1 GI:5974844
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2520)
 AUTHORS de la Monte, S. and Wands, J.R.
 TITLE Neural thread protein gene expression and detection of Alzheimer's
 disease
 JOURNAL Patent: US 5830670-A 50 03-NOV-1998;
 Location/Qualifiers
 1. 2520
 /organism="unknown"
 BASE COUNT 660 a 578 c 581 g 701 t
 ORIGIN

Query Match 3.0%; Score 18; DB 9; Length 2520;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 cccctgcatacccccgcg 247
 ||||||||||||||||
 Db 298 CCCCTGCATCACCCCTCG 315

RESULT 49
 AR072620
 LOCUS AR072620 2520 bp DNA PAT 28-AUG-2000
 DEFINITION Sequence 50 from patent US 5948634.
 ACCESSION AR072620
 VERSION AR072620.1 GI:9999384
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2520)
 AUTHORS de la Monte, S. and Wands, J.R.
 TITLE Neural thread protein gene expression and detection of Alzheimer's

disease
JOURNAL Patent: US 5948634-A 50 07-SEP-1999;
FEATURES Location/Qualifiers
source 1..2520
/organism="unknown"
BASE COUNT 660 a 578 c 581 g 701 t
ORIGIN

Query Match 3.0%; Score 18; DB 9; Length 2520;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 230 cccctgcacccccctg 247
|||||
DB 298 CCCCTGCATCACCCCTG 315

RESULT 50
AR073165
LOCUS AR073165 2520 bp DNA PAT 28-AUG-2000
DEFINITION Sequence 50 from patent US 5948634.
ACCESSION AR073165
VERSION AR073165.1 GI:9999928
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2520)
AUTHORS de la Monte,S. and Wands,J.R.
TITLE Neutral thread protein gene expression and detection of Alzheimer's
disease
JOURNAL Patent: US 5948634-A 50 07-SEP-1999;
FEATURES Location/Qualifiers
source 1..2520
/organism="unknown"
BASE COUNT 660 a 578 c 581 g 701 t
ORIGIN

Query Match 3.0%; Score 18; DB 9; Length 2520;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 230 cccctgcacccccctg 247
|||||
DB 298 CCCCTGCATCACCCCTG 315

Search completed: September 21, 2001, 21:10:19
Job time: 25839 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 13:47:10 : Search time 646.95 Seconds
(without alignments)
576.511 Million cell updates/sec

Title: US-09-138-735-1_COPY_1232_1825

Perfect score: 594
Sequence: 1 caggtacagcgtaacgagctt.....aaacgagctccgtagcgata 594

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 730101 seqs, 313950809 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N.Geneseq_0601.*

```
1: /cgnl_9/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /cgnl_9/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /cgnl_9/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /cgnl_9/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /cgnl_9/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /cgnl_9/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /cgnl_9/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /cgnl_9/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /cgnl_9/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /cgnl_9/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /cgnl_9/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /cgnl_9/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /cgnl_9/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /cgnl_9/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /cgnl_9/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /cgnl_9/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /cgnl_9/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /cgnl_9/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /cgnl_9/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /cgnl_9/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /cgnl_9/gcgdata/geneseq/geneseqn/NA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 594 | 100.0 | 3402 | 20 | AAH84092 |
| 2 | 492 | 82.8 | 3402 | 17 | AAT27310 |
| 3 | 22 | 3.7 | 22 | 17 | AAT27313 |
| 4 | 22 | 3.7 | 22 | 17 | AAT27313 |
| 5 | 19 | 3.2 | 11878 | 19 | AAV30199 |
| 6 | 19 | 3.2 | 11883 | 19 | AAV30198 |
| 7 | 18 | 3.0 | 18 | 17 | AAT27312 |
| 8 | 18 | 3.0 | 18 | 20 | AAH84094 |
| 9 | 18 | 3.0 | 454 | 21 | AAH09309 |
| 10 | 18 | 3.0 | 479 | 21 | AAH04783 |
| 11 | 18 | 3.0 | 866 | 20 | AAH39943 |

| | | | | | |
|----|----|-----|--------|----|-----------|
| 12 | 18 | 3.0 | 1302 | 21 | AAH59435 |
| 13 | 18 | 3.0 | 2520 | 17 | AAH07884 |
| 14 | 18 | 3.0 | 2520 | 17 | AAT27766 |
| 15 | 18 | 3.0 | 68750 | 21 | AAH35887 |
| 16 | 18 | 3.0 | 71989 | 21 | AAH29349 |
| 17 | 17 | 2.9 | 198 | 20 | AAH84320 |
| 18 | 17 | 2.9 | 353 | 21 | AAH09227 |
| 19 | 17 | 2.9 | 386 | 21 | AAH01025 |
| 20 | 17 | 2.9 | 513 | 21 | AAH38084 |
| 21 | 17 | 2.9 | 673 | 19 | AAH16884 |
| 22 | 17 | 2.9 | 673 | 20 | AAH26019 |
| 23 | 17 | 2.9 | 697 | 21 | AAH87504 |
| 24 | 17 | 2.9 | 697 | 21 | AAH33534 |
| 25 | 17 | 2.9 | 763 | 21 | AAH39218 |
| 26 | 17 | 2.9 | 800 | 21 | AAH9839 |
| 27 | 17 | 2.9 | 1148 | 21 | AAH42507 |
| 28 | 17 | 2.9 | 1273 | 21 | AAH4984 |
| 29 | 17 | 2.9 | 1338 | 19 | AAH64526 |
| 30 | 17 | 2.9 | 1338 | 19 | AAH44417 |
| 31 | 17 | 2.9 | 1338 | 20 | AAH19327 |
| 32 | 17 | 2.9 | 1338 | 20 | AAH19115 |
| 33 | 17 | 2.9 | 1338 | 20 | AAH21399 |
| 34 | 17 | 2.9 | 1647 | 21 | AAH79935 |
| 35 | 17 | 2.9 | 3337 | 17 | AAH34620 |
| 36 | 17 | 2.9 | 3337 | 20 | AAH15174 |
| 37 | 17 | 2.9 | 6171 | 21 | AAH74895 |
| 38 | 17 | 2.9 | 11212 | 21 | AAH58762 |
| 39 | 17 | 2.9 | 11672 | 21 | AAH01009 |
| 40 | 17 | 2.9 | 15611 | 21 | AAH01008 |
| 41 | 17 | 2.9 | 240825 | 22 | AAH24497 |
| 42 | 16 | 2.7 | 41 | 15 | AAH36842 |
| 43 | 16 | 2.7 | 222 | 15 | AAH073710 |
| 44 | 16 | 2.7 | 334 | 20 | AAH25170 |
| 45 | 16 | 2.7 | 371 | 21 | AAH21399 |
| 46 | 16 | 2.7 | 398 | 21 | AAH01286 |
| 47 | 16 | 2.7 | 413 | 21 | AAH79394 |
| 48 | 16 | 2.7 | 489 | 21 | AAH01726 |
| 49 | 16 | 2.7 | 541 | 19 | AAH76907 |
| 50 | 16 | 2.7 | 624 | 18 | AAH84039 |
| 51 | 16 | 2.7 | 651 | 18 | AAH15445 |
| 52 | 16 | 2.7 | 774 | 19 | AAH28667 |
| 53 | 16 | 2.7 | 774 | 20 | AAH216622 |
| 54 | 16 | 2.7 | 1185 | 19 | AAH05039 |
| 55 | 16 | 2.7 | 1185 | 20 | AAH08784 |
| 56 | 16 | 2.7 | 1185 | 21 | AAH71692 |
| 57 | 16 | 2.7 | 1234 | 21 | AAH35983 |
| 58 | 16 | 2.7 | 1314 | 19 | AAH00732 |
| 59 | 16 | 2.7 | 1314 | 19 | AAH09626 |
| 60 | 16 | 2.7 | 1314 | 20 | AAH07721 |
| 61 | 16 | 2.7 | 1411 | 14 | AAH041260 |
| 62 | 16 | 2.7 | 1411 | 15 | AAH072717 |
| 63 | 16 | 2.7 | 1669 | 20 | AAH25204 |
| 64 | 16 | 2.7 | 1864 | 20 | AAH24942 |
| 65 | 16 | 2.7 | 2063 | 22 | AAH84244 |
| 66 | 16 | 2.7 | 2214 | 19 | AAH22682 |
| 67 | 16 | 2.7 | 2274 | 21 | AAH47059 |
| 68 | 16 | 2.7 | 2310 | 20 | AAH21502 |
| 69 | 16 | 2.7 | 2610 | 21 | AAH90578 |
| 70 | 16 | 2.7 | 2886 | 21 | AAH251244 |
| 71 | 16 | 2.7 | 3045 | 20 | AAH21185 |
| 72 | 16 | 2.7 | 3045 | 20 | AAH21185 |
| 73 | 16 | 2.7 | 3331 | 19 | AAH22683 |
| 74 | 16 | 2.7 | 3331 | 20 | AAH238021 |
| 75 | 16 | 2.7 | 3331 | 22 | AAH39078 |
| 76 | 16 | 2.7 | 5117 | 21 | AAH06027 |
| 77 | 16 | 2.7 | 6242 | 21 | AAH99495 |
| 78 | 16 | 2.7 | 6545 | 17 | AAH27537 |
| 79 | 16 | 2.7 | 7732 | 20 | AAH77723 |
| 80 | 16 | 2.7 | 9243 | 21 | AAH81534 |
| 81 | 16 | 2.7 | 9551 | 20 | AAH22301 |
| 82 | 16 | 2.7 | 16510 | 21 | AAH21086 |
| 83 | 16 | 2.7 | 16510 | 21 | AAH34964 |
| 84 | 16 | 2.7 | 17634 | 21 | AAH21087 |

Human secreted pro
Neural thread prot
AD 16c human neut
Sorantium cellulos
Stealth virus nucl
Human secreted pro
Human secreted pro
Arabidopsis thalia
Human prostate can
Prostate, breast a
Human prostate can
Arabidopsis thalia
Human secreted pro
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
M. tuberculosis im
Mycobacterium tube
M. tuberculosis an
M. tuberculosis re
Human secreted pro
P. vivax ESP-1 blo
DNA encoding a sec
Human ORFX ORF450
DNA encoding a cyc
Escherichia coli p
Escherichia coli p
Human PG-3 gene.
Adaptor-primer. S
Reticulansposon pro
HIV-1 group O Isol
3' fragment of Abs
Human secreted pro
Eucalyptus grandis
Human secreted pro
S. glaucosens Hst
DNA encoding a Sta
Human gene fragmen
Ripening banana pu
Human gene express
Alcaligenes eutrop
A. eutrophus beta-
A. eutrophus BxTB
Arabidopsis thalia
Apepogillius terreu
A. terreus FAOD-L
Heat-resistant fru
Encodes repressor
Dialkylglycine dec
Mazze cinnamate-4-
Wild type A.eutrop
Signal transductio
New DNA sequence i
Zea mays DNA fragm
Saccharopolyspora
Human death induce
Murine TANGO 130 c
Neisseria meningit
Neisseria gonorrhoe
New DNA sequence i
Human MERT1 relate
DB6074 cDNA clone.
hudep-1 cDNA. Hom
Human cDNA encodin
Maltoze binding pr
pV2 vector. Synt
N. meningitidis pa
cDNA encoding a hu
Human low adenosin
Human adenosine re
Human low adenosin

| | | | | | | | | | | | |
|-------|----|-----|---------|----|-----------|--------------------|-------|------|----|-----------|----------------------|
| C 85 | 16 | 2.7 | 17634 | 21 | AAA34965 | Human adenosine re | C 158 | 1102 | 22 | AAE222774 | Human prostate can |
| C 86 | 16 | 2.7 | 21721 | 20 | AAx83427 | Human lipolysis st | C 159 | 1129 | 21 | AAc54524 | Arabidopsis thalia |
| C 87 | 16 | 2.7 | 22976 | 20 | AAx83426 | Genomic region con | C 160 | 1144 | 21 | AAc33490 | Arabidopsis thalia |
| C 88 | 16 | 2.7 | 23187 | 21 | AAx50273 | Human lipolysis st | C 161 | 1189 | 21 | AAE07514 | Fusarium venenatum |
| C 89 | 16 | 2.7 | 23187 | 22 | AAE2331 | Human leptin fragm | C 162 | 1200 | 22 | AAE71690 | Corynebacterium gl |
| C 90 | 16 | 2.7 | 39796 | 21 | AAc61681 | Nucleotide sequenc | C 163 | 1211 | 21 | AAc41977 | Arabidopsis thalia |
| C 91 | 16 | 2.7 | 47475 | 21 | AAAB1465 | N. meningitidis pa | C 164 | 1221 | 21 | AAE97039 | Arabidopsis thalia |
| C 92 | 16 | 2.7 | 102634 | 21 | AAAB1464 | N. meningitidis pa | C 165 | 1223 | 21 | AAc99054 | Human gene 75 DNA |
| C 93 | 16 | 2.7 | 349980 | 21 | AAE21608 | Neisseria meningit | C 166 | 1263 | 21 | AAc35255 | Human pancreatic c |
| C 94 | 16 | 2.7 | 349980 | 21 | AAE21609 | Neisseria meningit | C 167 | 1278 | 20 | AAx90824 | Plant retroelement |
| C 95 | 16 | 2.7 | 349980 | 21 | AAE21609 | Neisseria meningit | C 168 | 1281 | 20 | AAx90824 | DNA encoding human |
| C 96 | 16 | 2.7 | 349980 | 21 | AAE21612 | Neisseria meningit | C 169 | 1287 | 20 | AAx15146 | Coding region for |
| C 97 | 16 | 2.7 | 837096 | 21 | AAAB1489 | N. meningitidis pa | C 170 | 1294 | 21 | AAE13611 | Corynebacterium gl |
| C 98 | 16 | 2.7 | 1437668 | 21 | AAAB1490 | N. meningitidis B | C 171 | 1313 | 16 | AAQ91637 | Aspergillus oryzae |
| C 99 | 15 | 2.5 | 20 | 21 | AAAL4833 | PCR primer HG52.16 | C 172 | 1313 | 16 | AAQ91637 | Mouse sonic hedgeh |
| C 100 | 15 | 2.5 | 40 | 18 | AAE58819 | DNP DU labelled ol | C 173 | 1313 | 20 | AAx25620 | Mouse sonic hedgeh |
| C 101 | 15 | 2.5 | 44 | 18 | AAE58819 | Glycerol-based DNP | C 174 | 1313 | 20 | AAx25101 | Mouse sonic hedgeh |
| C 102 | 15 | 2.5 | 44 | 18 | AAE58825 | Glycerol DNP/DNP D | C 175 | 1313 | 20 | AAx07274 | Mouse sonic hedgeh |
| C 103 | 15 | 2.5 | 81 | 16 | AAE06202 | HIV-1 reverse tran | C 176 | 1313 | 20 | AAx16185 | Mouse shh hedgehog |
| C 104 | 15 | 2.5 | 168 | 21 | AAx44666 | Human secreted exp | C 177 | 1313 | 21 | AAx50449 | Mouse sonic hedgeh |
| C 105 | 15 | 2.5 | 196 | 21 | AAx41667 | Human secreted exp | C 178 | 1313 | 21 | AAx30277 | Mouse sonic hedgeh |
| C 106 | 15 | 2.5 | 295 | 16 | AAE21139 | Human gene signalu | C 179 | 1313 | 21 | AAx52260 | Mouse sonic hedgeh |
| C 107 | 15 | 2.5 | 297 | 21 | AAx7752 | CDNA encoding huma | C 180 | 1313 | 22 | AAE87077 | Mouse sonic hedgeh |
| C 108 | 15 | 2.5 | 300 | 20 | AAx14302 | Human gene express | C 181 | 1313 | 22 | AAE27016 | Mouse sonic hedgeh |
| C 109 | 15 | 2.5 | 300 | 20 | AAx12540 | Human gene express | C 182 | 1314 | 22 | AAx37897 | Mouse sonic hedgeh |
| C 110 | 15 | 2.5 | 300 | 21 | AAx00491 | Human colon cancer | C 183 | 1340 | 21 | AAx26964 | Human coenzyme A-u |
| C 111 | 15 | 2.5 | 325 | 20 | AAx86047 | Exon 1 of the huma | C 184 | 1362 | 21 | AAx50399 | Human angiotensin |
| C 112 | 15 | 2.5 | 325 | 21 | AAE60344 | Human neurotrophid | C 185 | 1362 | 21 | AAx35278 | Soybean retroelme |
| C 113 | 15 | 2.5 | 325 | 21 | AAE60670 | Human secreted pro | C 186 | 1368 | 21 | AAx53670 | Neisseria gonorrhoe |
| C 114 | 15 | 2.5 | 328 | 21 | AAE12718 | Human secreted pro | C 187 | 1368 | 21 | AAE53672 | Neisseria meningit |
| C 115 | 15 | 2.5 | 362 | 22 | AAE93718 | CDNA encoding SFR | C 188 | 1368 | 22 | AAE72256 | Corynebacterium gl |
| C 116 | 15 | 2.5 | 384 | 22 | AAE65140 | Novel human polyu | C 189 | 1368 | 22 | AAE72257 | Corynebacterium gl |
| C 117 | 15 | 2.5 | 397 | 19 | AAx44835 | Novel HMSCX46r rel | C 190 | 1374 | 21 | AAE31258 | Arabidopsis thalia |
| C 118 | 15 | 2.5 | 397 | 22 | AAE65139 | Novel human polyu | C 191 | 1376 | 21 | AAc33903 | Arabidopsis thalia |
| C 119 | 15 | 2.5 | 399 | 22 | AAE64930 | Novel human polyu | C 192 | 1379 | 21 | AAE297132 | Human secreted pro |
| C 120 | 15 | 2.5 | 403 | 22 | AAE66880 | Nucleotide encodin | C 193 | 1381 | 20 | AAx84482 | Human secreted pro |
| C 121 | 15 | 2.5 | 426 | 20 | AAx8739 | Human secreted exp | C 194 | 1390 | 22 | AAE60982 | P. putida KT240-a |
| C 122 | 15 | 2.5 | 430 | 21 | AAx43600 | Polynucleotide seq | C 195 | 1396 | 22 | AAc33740 | Arabidopsis thalia |
| C 123 | 15 | 2.5 | 462 | 20 | AAx20850 | Porcine BAC-PIGF2- | C 196 | 1423 | 17 | AAE72257 | CDNA encoding cell |
| C 124 | 15 | 2.5 | 500 | 21 | AAx65460 | Eucalyptus grandis | C 197 | 1437 | 20 | AAx84973 | Human secreted pro |
| C 125 | 15 | 2.5 | 574 | 21 | AAE65056 | Human prostate can | C 198 | 1448 | 20 | AAx42009 | Human endometrium |
| C 126 | 15 | 2.5 | 603 | 22 | AAE22775 | Human prostate can | C 199 | 1486 | 21 | AAc38414 | Arabidopsis thalia |
| C 127 | 15 | 2.5 | 628 | 22 | AAE22779 | Arabidopsis thalia | C 200 | 1498 | 21 | AAc45265 | Arabidopsis thalia |
| C 128 | 15 | 2.5 | 631 | 21 | AAc54906 | Arabidopsis thalia | C 201 | 1572 | 20 | AAx02671 | T. versicolor lac |
| C 129 | 15 | 2.5 | 633 | 21 | AAc65696 | Arabidopsis thalia | C 202 | 1572 | 20 | AAx02672 | T. versicolor lac |
| C 130 | 15 | 2.5 | 633 | 21 | AAc65696 | Arabidopsis thalia | C 203 | 1591 | 16 | AAQ97586 | CDNA encoding prote |
| C 131 | 15 | 2.5 | 637 | 21 | AAc31309 | Arabidopsis thalia | C 204 | 1631 | 21 | AAQ41461 | Myobacterium para |
| C 132 | 15 | 2.5 | 658 | 21 | AAc37714 | Arabidopsis thalia | C 205 | 1631 | 21 | AAE98141 | Human signal pepi |
| C 133 | 15 | 2.5 | 659 | 21 | AAAB1591 | N. meningitidis pa | C 206 | 1632 | 14 | AAQ42976 | Carlinomycin 4-O-m |
| C 134 | 15 | 2.5 | 672 | 21 | AAc52269 | Human colon cancer | C 207 | 1667 | 21 | AAE51207 | Rat lost in leukae |
| C 135 | 15 | 2.5 | 674 | 21 | AAAB02040 | Human prostate can | C 208 | 1715 | 16 | AAE04588 | Rat Vhh-1 cDNA. R |
| C 136 | 15 | 2.5 | 675 | 22 | AAE22781 | Human secreted can | C 209 | 1738 | 21 | AAx51206 | Human lost in leukae |
| C 137 | 15 | 2.5 | 686 | 21 | AAc60054 | Human secreted pro | C 210 | 1755 | 20 | AAE06232 | Human secreted pro |
| C 138 | 15 | 2.5 | 791 | 21 | AAc54902 | Arabidopsis thalia | C 211 | 1779 | 20 | AAE12073 | Neisseria meningit |
| C 139 | 15 | 2.5 | 792 | 21 | AAc52503 | Arabidopsis thalia | C 212 | 1779 | 20 | AAE12074 | Neisseria meningit |
| C 140 | 15 | 2.5 | 799 | 20 | AAx16224 | Human gene express | C 213 | 1779 | 20 | AAE12075 | Neisseria gonorrhoe |
| C 141 | 15 | 2.5 | 800 | 19 | AAx63926 | Myobacterium tube | C 214 | 1779 | 21 | AAE53673 | Neisseria gonorrhoe |
| C 142 | 15 | 2.5 | 800 | 20 | AAx81033 | Nucleotide sequenc | C 215 | 1779 | 21 | AAE53674 | Neisseria meningit |
| C 143 | 15 | 2.5 | 810 | 21 | AAx39970 | Murine TANGO 188 c | C 216 | 1779 | 21 | AAE53675 | Neisseria meningit |
| C 144 | 15 | 2.5 | 837 | 21 | AAx30334 | Nucleotide sequenc | C 217 | 1788 | 22 | AAE32651 | Human cDNA encodin |
| C 145 | 15 | 2.5 | 838 | 16 | AAQ90275 | Human beta-kinesin | C 218 | 1827 | 21 | AAE69692 | Novel Human protei |
| C 146 | 15 | 2.5 | 841 | 22 | AAE75381 | Ralstonia eutropha | C 219 | 1831 | 20 | AAE26880 | DNA encoding a pro |
| C 147 | 15 | 2.5 | 846 | 20 | AAE25129 | Soybean chalcone 1 | C 220 | 1831 | 21 | AAx09889 | Human OCTN2 cDNA s |
| C 148 | 15 | 2.5 | 846 | 21 | AAx38424 | Sequence of region | C 221 | 1858 | 22 | AAE30190 | Clone 11696905-0-4 |
| C 149 | 15 | 2.5 | 876 | 8 | AAx70094 | DNA encoding a hum | C 222 | 1858 | 20 | AAx85599 | Human angiotensin |
| C 150 | 15 | 2.5 | 901 | 20 | AAx86270 | Clone HTAS03 codi | C 223 | 1862 | 22 | AAE02606 | Yeast essential ge |
| C 151 | 15 | 2.5 | 921 | 21 | AAc45812 | Arabidopsis thalia | C 224 | 1863 | 21 | AAx28268 | Human TIE ligand N |
| C 152 | 15 | 2.5 | 948 | 22 | AAc90012 | Arabidopsis thalia | C 225 | 1869 | 20 | AAx33970 | Human TIE ligand N |
| C 153 | 15 | 2.5 | 1005 | 21 | AAc51280 | Arabidopsis thalia | C 226 | 1869 | 21 | AAx60537 | Human TIE ligand N |
| C 154 | 15 | 2.5 | 1014 | 21 | AAc34499 | As protein coupled | C 227 | 1869 | 21 | AAx50126 | Human angiotensin |
| C 155 | 15 | 2.5 | 1029 | 17 | AAE72718 | Aspergillus oryzae | C 228 | 1873 | 22 | AAE97388 | Human neovasculari |
| C 156 | 15 | 2.5 | 1065 | 21 | AAE14557 | Nucleotide encodin | C 229 | 1876 | 12 | AAE27220 | Clone 1 for trunca |
| C 157 | 15 | 2.5 | 1078 | 20 | AAx08745 | | C 230 | 1876 | 12 | AAQ14498 | |

| | | | | | | | | | | | | | |
|-------|----|-----|------|----|-----------|--------------------|-------|----|-----|--------|----|-----------|--------------------|
| 231 | 15 | 2.5 | 1876 | 12 | AA014499 | Clone 2 for trunca | c 304 | 15 | 2.5 | 9706 | 21 | AAA65342 | NDO related comple |
| c 232 | 15 | 2.5 | 1876 | 20 | AA015144 | cDNA encoding huma | c 305 | 15 | 2.5 | 9840 | 22 | AA065190 | NDO related comple |
| c 233 | 15 | 2.5 | 1879 | 21 | AA045850 | Human muscle angio | c 306 | 15 | 2.5 | 9894 | 22 | AA085131 | S. avermitilis 10 |
| c 234 | 15 | 2.5 | 1879 | 22 | AA029701 | Human angiotensin | c 307 | 15 | 2.5 | 10427 | 21 | AA0236325 | Mechanical stress |
| c 235 | 15 | 2.5 | 1894 | 20 | AA087858 | Human fibrinogen d | c 308 | 15 | 2.5 | 10897 | 17 | AA0109187 | Mutp putative onco |
| c 236 | 15 | 2.5 | 1911 | 21 | AA022396 | Human secreted pro | c 309 | 15 | 2.5 | 10898 | 21 | AA052462 | Human MN gene. Ho |
| c 237 | 15 | 2.5 | 1913 | 22 | AA0297138 | Human gene 75 DNA | c 310 | 15 | 2.5 | 10898 | 21 | AA016543 | Human MN complete |
| c 238 | 15 | 2.5 | 1921 | 21 | AA044835 | RP55-like protein | c 311 | 15 | 2.5 | 11279 | 21 | AA038389 | Pseudomonas sp. WP |
| c 239 | 15 | 2.5 | 1931 | 21 | AA0297093 | Human secreted pro | c 312 | 15 | 2.5 | 11871 | 20 | AA013108 | Enterococcus faeca |
| c 240 | 15 | 2.5 | 1955 | 21 | AA0299588 | cDNA encoding a hu | c 313 | 15 | 2.5 | 12286 | 21 | AA035261 | Plant genetic retr |
| c 241 | 15 | 2.5 | 1973 | 21 | AA060998 | Human vesicle asso | c 314 | 15 | 2.5 | 12571 | 21 | AA035272 | Soybean retroleme |
| c 242 | 15 | 2.5 | 2044 | 17 | AA013167 | Rat interleukin-1 | c 315 | 15 | 2.5 | 12808 | 21 | AA065347 | NDO related comple |
| c 243 | 15 | 2.5 | 2044 | 18 | AA084010 | DNA encoding a Sta | c 316 | 15 | 2.5 | 12886 | 21 | AA059888 | Human patched-like |
| c 244 | 15 | 2.5 | 2051 | 20 | AA028069 | Heparin sulphate 6 | c 317 | 15 | 2.5 | 13518 | 20 | AA020563 | Human patched-like |
| c 245 | 15 | 2.5 | 2051 | 21 | AA0299186 | Human HS6ST1 cDNA. | c 318 | 15 | 2.5 | 13574 | 21 | AA081529 | polynucleotide seq |
| c 246 | 15 | 2.5 | 2084 | 21 | AA052492 | Human secreted pro | c 319 | 15 | 2.5 | 16020 | 21 | AA035283 | N. meningitidis pa |
| c 247 | 15 | 2.5 | 2102 | 21 | AA039969 | Human secreted pro | c 320 | 15 | 2.5 | 24417 | 18 | AA0197221 | Streptomyces nogal |
| c 248 | 15 | 2.5 | 2104 | 22 | AA087120 | Murine TANGO 188 c | c 321 | 15 | 2.5 | 25871 | 21 | AA098888 | pseudomonas aerugi |
| c 249 | 15 | 2.5 | 2160 | 20 | AA086635 | Nucleotide sequenc | c 322 | 15 | 2.5 | 42000 | 21 | AA063349 | Human genomic OCTN |
| c 250 | 15 | 2.5 | 2187 | 21 | AA029701 | Wild-type human c- | c 323 | 15 | 2.5 | 44576 | 21 | AA061522 | Streptomyces globi |
| c 251 | 15 | 2.5 | 2200 | 22 | AA027713 | Human transport pr | c 324 | 15 | 2.5 | 50341 | 19 | AA022674 | Cosmid CVO14 conta |
| c 252 | 15 | 2.5 | 2221 | 21 | AA030370 | Partial nucleotide | c 325 | 15 | 2.5 | 50341 | 19 | AA022674 | DNA sequence of a |
| c 253 | 15 | 2.5 | 2249 | 21 | AA037055 | Human PKO1374 (UNQ | c 326 | 15 | 2.5 | 50341 | 21 | AA039519 | L5 shuttle phasmid |
| c 254 | 15 | 2.5 | 2249 | 22 | AA054285 | DNA encoding prote | c 327 | 15 | 2.5 | 52297 | 16 | AA051411 | Streptococcus olea |
| c 255 | 15 | 2.5 | 2263 | 20 | AA080628 | Kidney injury asso | c 328 | 15 | 2.5 | 52298 | 14 | AA047357 | Mycobacteriophage |
| c 256 | 15 | 2.5 | 2271 | 11 | AA006844 | Amylase gene from | c 329 | 15 | 2.5 | 58857 | 21 | AA058471 | L5 mycobacterioph |
| c 257 | 15 | 2.5 | 2276 | 18 | AA086161 | PHA depolymerase g | c 330 | 15 | 2.5 | 63164 | 21 | AA063348 | Nucleotide sequenc |
| c 258 | 15 | 2.5 | 2291 | 9 | AA080309 | Entire amyase gen | c 331 | 15 | 2.5 | 78845 | 21 | AA081463 | Streptomyces globi |
| c 259 | 15 | 2.5 | 2311 | 21 | AA018225 | Lung cancer associ | c 332 | 15 | 2.5 | 109973 | 21 | AA081463 | N. meningitidis pa |
| c 260 | 15 | 2.5 | 2324 | 21 | AA015288 | cDNA encoding prot | c 333 | 15 | 2.5 | 109973 | 21 | AA022298 | BAC containing rep |
| c 261 | 15 | 2.5 | 2332 | 21 | AA072398 | Human nucleic acid | c 334 | 15 | 2.5 | 235033 | 21 | AA021613 | Neisseria meningit |
| c 262 | 15 | 2.5 | 2420 | 21 | AA046489 | PKA substrate, Csk | c 335 | 15 | 2.5 | 237326 | 19 | AA057926 | Hereditary haemoch |
| c 263 | 15 | 2.5 | 2501 | 21 | AA052499 | 2.5 kb human MN ge | c 336 | 15 | 2.5 | 349980 | 21 | AA021608 | Neisseria meningit |
| c 264 | 15 | 2.5 | 2501 | 21 | AA016580 | MN genomic region | c 337 | 15 | 2.5 | 837096 | 20 | AA021608 | N. meningitidis pa |
| c 265 | 15 | 2.5 | 2507 | 21 | AA052365 | NSQ gene-17 assoc | c 338 | 15 | 2.5 | 837096 | 20 | AA081489 | PCR primer used to |
| c 266 | 15 | 2.5 | 2510 | 20 | AA041256 | Human normal ovary | c 339 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 267 | 15 | 2.5 | 2510 | 20 | AA041256 | Human breast tumou | c 340 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 268 | 15 | 2.5 | 2524 | 21 | AA077206 | Human ORFX ORF2761 | c 341 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 269 | 15 | 2.5 | 2623 | 21 | AA075996 | Human ORFX ORF1551 | c 342 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 270 | 15 | 2.5 | 2635 | 20 | AA084521 | Human secreted pro | c 343 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 271 | 15 | 2.5 | 2646 | 16 | AA082405 | Rice mature pullul | c 344 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 272 | 15 | 2.5 | 2762 | 9 | AA080976 | Sequence of 2 cont | c 345 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 273 | 15 | 2.5 | 2768 | 12 | AA010502 | Pulihydroxybutyrat | c 346 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 274 | 15 | 2.5 | 2982 | 16 | AA024000 | Pullulanase expres | c 347 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 275 | 15 | 2.5 | 2988 | 16 | AA023399 | Rice mature pullul | c 348 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 276 | 15 | 2.5 | 2999 | 21 | AA078119 | Human cancer assoc | c 349 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 277 | 15 | 2.5 | 3035 | 21 | AA063462 | Streptomyces globi | c 350 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 278 | 15 | 2.5 | 3103 | 21 | AA050708 | Nucleotide sequenc | c 351 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 279 | 15 | 2.5 | 3105 | 18 | AA074840 | Staphylococcus aur | c 352 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 280 | 15 | 2.5 | 3105 | 18 | AA074840 | Human FcH540 gene | c 353 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 281 | 15 | 2.5 | 3111 | 22 | AA026660 | Human Smad7 nucleo | c 354 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 282 | 15 | 2.5 | 3141 | 14 | AA049903 | Nicotiana tabacum | c 355 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 283 | 15 | 2.5 | 3218 | 22 | AA063440 | Human ADAMTS-7 CDN | c 356 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 284 | 15 | 2.5 | 3284 | 20 | AA002677 | T. versicolor lacC | c 357 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 285 | 15 | 2.5 | 3300 | 18 | AA062589 | Thermotoga neopoli | c 358 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 286 | 15 | 2.5 | 3309 | 21 | AA021886 | Human breast and o | c 359 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 287 | 15 | 2.5 | 3337 | 17 | AA091089 | Human nerve growth | c 360 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 288 | 15 | 2.5 | 3517 | 10 | AA031287 | Fragment of Pseudo | c 361 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 289 | 15 | 2.5 | 3532 | 21 | AA052530 | Rabbit poly-immuno | c 362 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 290 | 15 | 2.5 | 3532 | 21 | AA052530 | Human MN 5' UTR fr | c 363 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 291 | 15 | 2.5 | 3821 | 21 | AA016611 | MN promoter region | c 364 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 292 | 15 | 2.5 | 3821 | 21 | AA056609 | Human Pgf-2/NT-3.5 | c 365 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 293 | 15 | 2.5 | 3821 | 21 | AA056609 | Human nerve growth | c 366 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 294 | 15 | 2.5 | 3981 | 19 | AA040381 | Strawberry structu | c 367 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 295 | 15 | 2.5 | 4030 | 20 | AA031717 | Human patched-2 co | c 368 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 296 | 15 | 2.5 | 4296 | 20 | AA033436 | Rat TAO2 kinase en | c 369 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 297 | 15 | 2.5 | 4751 | 20 | AA029496 | Human heart tissue | c 370 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 298 | 15 | 2.5 | 4879 | 18 | AA074400 | Staphylococcus aur | c 371 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 299 | 15 | 2.5 | 4980 | 14 | AA036660 | Sequence of the po | c 372 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 300 | 15 | 2.5 | 5342 | 21 | AA018986 | polyhydroxyalkano | c 373 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 301 | 15 | 2.5 | 5342 | 21 | AA000374 | Rat smooth muscle- | c 374 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 302 | 15 | 2.5 | 5631 | 20 | AA098274 | Human tumour suppr | c 375 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 303 | 15 | 2.5 | 6505 | 21 | AA075588 | Human ORFX ORF1143 | c 376 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |
| c 303 | 15 | 2.5 | 6570 | 20 | AA020540 | Polynucleotide seq | c 376 | 15 | 2.5 | 837096 | 20 | AA081489 | Human cytochrome p |

| | | | | | | |
|-----|----|-----|-----|----|-----------|---------------------|
| 377 | 14 | 2.4 | 282 | 14 | AA042653 | PT-NANB virus lamb |
| 378 | 14 | 2.4 | 294 | 14 | AA05481 | Human secreted pro |
| 379 | 14 | 2.4 | 295 | 15 | AA073711 | Retinopospon pro |
| 380 | 14 | 2.4 | 295 | 21 | AA06337 | Human secreted pro |
| 381 | 14 | 2.4 | 297 | 17 | AA09205 | Virulence factor s |
| 382 | 14 | 2.4 | 298 | 17 | AA09207 | Virulence factor s |
| 383 | 14 | 2.4 | 300 | 20 | AA012673 | Human gene express |
| 384 | 14 | 2.4 | 300 | 21 | AA00198 | Human colon cancer |
| 385 | 14 | 2.4 | 300 | 21 | AA00442 | Human colon cancer |
| 386 | 14 | 2.4 | 300 | 21 | AA01374 | Human colon cancer |
| 387 | 14 | 2.4 | 303 | 16 | AA018914 | DNA encoding spide |
| 388 | 14 | 2.4 | 304 | 21 | AA03031 | Human secreted pro |
| 389 | 14 | 2.4 | 304 | 21 | AA02955 | Human 5' EST Isola |
| 390 | 14 | 2.4 | 312 | 21 | AA064583 | Nucleotide sequenc |
| 391 | 14 | 2.4 | 313 | 16 | AA019149 | Human gene signatu |
| 392 | 14 | 2.4 | 314 | 21 | AA023276 | Human secreted pro |
| 393 | 14 | 2.4 | 316 | 21 | AA006622 | Human secreted pro |
| 394 | 14 | 2.4 | 316 | 21 | AA042333 | Human 5' EST Isola |
| 395 | 14 | 2.4 | 322 | 21 | AA04961 | Human ORFX ORF516 |
| 396 | 14 | 2.4 | 324 | 21 | AA071620 | N. meningitidis BA |
| 397 | 14 | 2.4 | 324 | 21 | AA053434 | Neisseria meningit |
| 398 | 14 | 2.4 | 324 | 21 | AA053435 | Neisseria meningit |
| 399 | 14 | 2.4 | 326 | 19 | AA019165 | Human XAG growth f |
| 400 | 14 | 2.4 | 326 | 19 | AA063340 | HuXAG-2 related cd |
| 401 | 14 | 2.4 | 330 | 20 | AA089310 | EST clone CM423. |
| 402 | 14 | 2.4 | 332 | 21 | AA08918 | Fusarium venenatum |
| 403 | 14 | 2.4 | 332 | 21 | AA08918 | Singlu nucleotide |
| 404 | 14 | 2.4 | 338 | 21 | AA014174 | Human secreted pro |
| 405 | 14 | 2.4 | 341 | 21 | AA011086 | Fusarium venenatum |
| 406 | 14 | 2.4 | 343 | 20 | AA028036 | Partial sequence o |
| 407 | 14 | 2.4 | 346 | 21 | AA015981 | Human secreted pro |
| 408 | 14 | 2.4 | 353 | 20 | AA086600 | EST clone BE101. |
| 409 | 14 | 2.4 | 354 | 22 | AA066171 | Novel human polynu |
| 410 | 14 | 2.4 | 363 | 21 | AA054189 | Neisseria meningit |
| 411 | 14 | 2.4 | 363 | 21 | AA054190 | Neisseria meningit |
| 412 | 14 | 2.4 | 375 | 21 | AA041703 | Arabidopsis thalia |
| 413 | 14 | 2.4 | 378 | 22 | AA064420 | M tuberculosis Rv3 |
| 414 | 14 | 2.4 | 381 | 14 | AA060707 | Human Brain Expres |
| 415 | 14 | 2.4 | 386 | 9 | AA080184 | DNA encoding modif |
| 416 | 14 | 2.4 | 387 | 9 | AA080192 | DNA encoding biosy |
| 417 | 14 | 2.4 | 391 | 22 | AA064566 | Novel human polynu |
| 418 | 14 | 2.4 | 394 | 22 | AA05032 | Human secreted pro |
| 419 | 14 | 2.4 | 395 | 21 | AA023666 | DNA encoding light |
| 420 | 14 | 2.4 | 396 | 21 | AA059695 | Human secreted pro |
| 421 | 14 | 2.4 | 396 | 21 | AA02889 | Human secreted pro |
| 422 | 14 | 2.4 | 397 | 20 | AA087051 | EST clone B197. H |
| 423 | 14 | 2.4 | 398 | 17 | AA061228 | Human Brain Expres |
| 424 | 14 | 2.4 | 399 | 17 | AA04932 | Partial sequence o |
| 425 | 14 | 2.4 | 399 | 17 | AA04936 | Partial sequence o |
| 426 | 14 | 2.4 | 399 | 17 | AA04937 | Partial sequence o |
| 427 | 14 | 2.4 | 399 | 17 | AA04939 | Partial sequence o |
| 428 | 14 | 2.4 | 400 | 10 | AA091652 | FC binding region |
| 429 | 14 | 2.4 | 400 | 18 | AA078295 | staphylococcus aur |
| 430 | 14 | 2.4 | 403 | 18 | AA075454 | staphylococcus aur |
| 431 | 14 | 2.4 | 403 | 22 | AA064800 | Novel human polynu |
| 432 | 14 | 2.4 | 405 | 20 | AA001709 | Human anti-angioge |
| 433 | 14 | 2.4 | 408 | 20 | AA090910 | Nucleotide sequenc |
| 434 | 14 | 2.4 | 408 | 20 | AA090625 | Nucleotide sequenc |
| 435 | 14 | 2.4 | 409 | 21 | AA074651 | Human ORFX ORF206 |
| 436 | 14 | 2.4 | 409 | 22 | AA067423 | Novel human polynu |
| 437 | 14 | 2.4 | 412 | 19 | AA028851 | Targeting vector T |
| 438 | 14 | 2.4 | 420 | 21 | AA026868 | Human secreted pro |
| 439 | 14 | 2.4 | 421 | 14 | AA038597 | Human glucose regu |
| 440 | 14 | 2.4 | 421 | 20 | AA088382 | EST clone DD285. |
| 441 | 14 | 2.4 | 427 | 22 | AA068261 | Human lung tumour |
| 442 | 14 | 2.4 | 433 | 21 | AA065825 | E. coli proliferat |
| 443 | 14 | 2.4 | 435 | 11 | AA005642 | Clone hp53 encodin |
| 444 | 14 | 2.4 | 436 | 20 | AA089723 | EST clone CT738. |
| 445 | 14 | 2.4 | 438 | 21 | AA048195 | Arabidopsis thalia |
| 446 | 14 | 2.4 | 438 | 21 | AA066074 | E. coli proliferat |
| 447 | 14 | 2.4 | 441 | 21 | AA067106 | Eucalyptus grandis |
| 448 | 14 | 2.4 | 443 | 21 | AA056803 | Eucalyptus grandis |
| 449 | 14 | 2.4 | 446 | 21 | AA020933 | Human secreted pro |
| 450 | 14 | 2.4 | 449 | 20 | AA089588 | EST clone CT729. |
| 451 | 14 | 2.4 | 458 | 21 | AA001824 | Human secreted pro |
| 452 | 14 | 2.4 | 461 | 21 | AA006277 | Human secreted pro |
| 453 | 14 | 2.4 | 464 | 21 | AA067550 | Pinus radiata sucr |
| 454 | 14 | 2.4 | 466 | 21 | AA053503 | Zea mays DNA fragm |
| 455 | 14 | 2.4 | 466 | 22 | AA065887 | Novel human polynu |
| 456 | 14 | 2.4 | 466 | 22 | AA065887 | Novel human polynu |
| 457 | 14 | 2.4 | 473 | 21 | AA009968 | Human secreted pro |
| 458 | 14 | 2.4 | 473 | 21 | AA021638 | Human secreted pro |
| 459 | 14 | 2.4 | 477 | 21 | AA076187 | Human ORFX ORF1742 |
| 460 | 14 | 2.4 | 478 | 21 | AA006307 | Human secreted pro |
| 461 | 14 | 2.4 | 481 | 13 | AA029869 | Phenomone receptor |
| 462 | 14 | 2.4 | 481 | 18 | AA075041 | Staphylococcus aur |
| 463 | 14 | 2.4 | 483 | 19 | AA060043 | Nucleic acid E69r |
| 464 | 14 | 2.4 | 488 | 21 | AA039752 | Zea mays DNA fragm |
| 465 | 14 | 2.4 | 489 | 20 | AA091716 | Porphyromonas ging |
| 466 | 14 | 2.4 | 492 | 21 | AA076086 | Human ORFX ORF1621 |
| 467 | 14 | 2.4 | 492 | 22 | AA054745 | Nucleotide sequenc |
| 468 | 14 | 2.4 | 500 | 21 | AA040965 | Zea mays DNA fragm |
| 469 | 14 | 2.4 | 501 | 17 | AA070739 | Immunogen DNA from |
| 470 | 14 | 2.4 | 505 | 19 | AA019167 | Human XAG growth f |
| 471 | 14 | 2.4 | 505 | 22 | AA063342 | HuXAG-2 related cd |
| 472 | 14 | 2.4 | 507 | 21 | AA051616 | Arabidopsis thalia |
| 473 | 14 | 2.4 | 508 | 21 | AA053128 | Arabidopsis thalia |
| 474 | 14 | 2.4 | 510 | 20 | AA091589 | Porphyromonas ging |
| 475 | 14 | 2.4 | 512 | 21 | AA052491 | Human MN gene intr |
| 476 | 14 | 2.4 | 512 | 21 | AA016572 | Human MN intion 7 |
| 477 | 14 | 2.4 | 513 | 17 | AA042239 | HIV-1 group O stra |
| 478 | 14 | 2.4 | 513 | 21 | AA070693 | Fusarium venenatum |
| 479 | 14 | 2.4 | 513 | 21 | AA040819 | Arabidopsis thalia |
| 480 | 14 | 2.4 | 516 | 20 | AA028678 | Clone HP01766 enco |
| 481 | 14 | 2.4 | 516 | 22 | AA045139 | Murine TANGO 273 O |
| 482 | 14 | 2.4 | 518 | 21 | AA014992 | Trichoderma reesei |
| 483 | 14 | 2.4 | 520 | 20 | AA091014 | Human secreted pro |
| 484 | 14 | 2.4 | 522 | 20 | AA026098 | S. aureus phosphor |
| 485 | 14 | 2.4 | 523 | 21 | AA034976 | Mouse F12Z2 (fnh1b |
| 486 | 14 | 2.4 | 524 | 21 | AA015310 | Trichoderma reesei |
| 487 | 14 | 2.4 | 525 | 20 | AA026097 | S. aureus phosphor |
| 488 | 14 | 2.4 | 525 | 21 | AA011017 | Fusarium venenatum |
| 489 | 14 | 2.4 | 527 | 22 | AA044783 | Testis Enhanced Ge |
| 490 | 14 | 2.4 | 534 | 14 | AA045088 | Human glucokinase |
| 491 | 14 | 2.4 | 534 | 14 | AA041505 | Arabidopsis thalia |
| 492 | 14 | 2.4 | 535 | 19 | AA022070 | DNA encoding a krl |
| 493 | 14 | 2.4 | 535 | 21 | AA090474 | White shrimp multi |
| 494 | 14 | 2.4 | 536 | 19 | AA022071 | DNA encoding a krl |
| 495 | 14 | 2.4 | 536 | 21 | AA090471 | White shrimp multi |
| 496 | 14 | 2.4 | 545 | 21 | AA0909347 | Fusarium venenatum |
| 497 | 14 | 2.4 | 549 | 16 | AA084578 | Pseudomonas aerugi |
| 498 | 14 | 2.4 | 550 | 21 | AA093988 | Cat flea hindgut a |
| 499 | 14 | 2.4 | 556 | 21 | AA041937 | Human secreted exp |
| 500 | 14 | 2.4 | 573 | 21 | AA080762 | Fusarium venenatum |
| 501 | 14 | 2.4 | 575 | 21 | AA078221 | Human cancer assoc |
| 502 | 14 | 2.4 | 579 | 20 | AA001708 | Human anti-angioge |
| 503 | 14 | 2.4 | 597 | 21 | AA016408 | Human prostate can |
| 504 | 14 | 2.4 | 597 | 21 | AA076643 | Human cancer assoc |
| 505 | 14 | 2.4 | 597 | 21 | AA030451 | Human brain cDNA c |
| 506 | 14 | 2.4 | 597 | 21 | AA015654 | HTLV p24 GI fragme |
| 507 | 14 | 2.4 | 599 | 17 | AA0708608 | Protoplasming memb |
| 508 | 14 | 2.4 | 599 | 19 | AA022065 | DNA encoding a krl |
| 509 | 14 | 2.4 | 623 | 21 | AA090470 | White shrimp multi |
| 510 | 14 | 2.4 | 623 | 21 | AA087734 | Human secreted pro |
| 511 | 14 | 2.4 | 623 | 21 | AA033194 | Arabidopsis thalia |
| 512 | 14 | 2.4 | 623 | 22 | AA064016 | Arabidopsis thalia |
| 513 | 14 | 2.4 | 624 | 8 | AA071272 | cDNA encoding huma |
| 514 | 14 | 2.4 | 624 | 21 | AA057319 | Sequence of adult |
| 515 | 14 | 2.4 | 626 | 21 | AA041611 | Pinus radiata thian |
| 516 | 14 | 2.4 | 631 | 21 | AA033898 | Arabidopsis thalia |
| 517 | 14 | 2.4 | 633 | 19 | AA022068 | DNA encoding a krl |
| 518 | 14 | 2.4 | 633 | 21 | AA090472 | White shrimp multi |
| 519 | 14 | 2.4 | 634 | 21 | AA048206 | Arabidopsis thalia |
| 520 | 14 | 2.4 | 636 | 21 | AA065944 | E. coli proliferat |
| 521 | 14 | 2.4 | 637 | 21 | AA012391 | Aspergillus oryzae |
| 522 | 14 | 2.4 | 637 | 21 | AA012850 | Aspergillus oryzae |
| 523 | 14 | 2.4 | 637 | 21 | AA012850 | Human secreted pro |

| | | | | | | | | | | | | |
|-------|----|------|----|----------|---------------------|-------|----|-----|------|----|-----------|----------------------|
| C 669 | 14 | 1147 | 21 | AA664274 | Human membrane-ass | 742 | 14 | 2.4 | 1392 | 18 | AA799938 | Bovine dipeptidyl |
| C 670 | 14 | 1149 | 21 | AA51824 | Human melancortin | C 743 | 14 | 2.4 | 1394 | 9 | AA82030 | Synthetic Hydroxym |
| C 671 | 14 | 1149 | 21 | AA51825 | Human melancortin | C 744 | 14 | 2.4 | 1394 | 9 | AA82030 | Synthetic human 3-hy |
| C 672 | 14 | 1149 | 21 | AA51826 | Human melancortin | C 745 | 14 | 2.4 | 1401 | 19 | AA734002 | S. penicellus dnrx |
| C 673 | 14 | 1149 | 21 | AA51827 | Human melancortin | C 746 | 14 | 2.4 | 1414 | 20 | AA728679 | Clone HP01766 enco |
| C 674 | 14 | 1149 | 21 | AA51828 | Human melancortin | C 747 | 14 | 2.4 | 1417 | 18 | AA793887 | CDNA encoding a no |
| C 675 | 14 | 1149 | 21 | AA51829 | Human melancortin | C 748 | 14 | 2.4 | 1417 | 18 | AA787401 | AdA/AD3LP sequence |
| C 676 | 14 | 1149 | 21 | AA51830 | Human melancortin | C 749 | 14 | 2.4 | 1417 | 20 | AA762129 | Human phosphorilas |
| C 677 | 14 | 1149 | 21 | AA51832 | Human melancortin | C 750 | 14 | 2.4 | 1419 | 21 | AA738239 | Novel human phosph |
| C 678 | 14 | 1149 | 21 | AA51834 | Human melancortin | C 751 | 14 | 2.4 | 1420 | 21 | AA742087 | Arbidopsin thalia |
| C 679 | 14 | 1149 | 21 | AA51836 | Human melancortin | C 752 | 14 | 2.4 | 1423 | 19 | AA719156 | Human XAG growth f |
| C 680 | 14 | 1149 | 21 | AA51838 | Human melancortin | C 753 | 14 | 2.4 | 1423 | 22 | AA763315 | Human huxAG-2/CSG |
| C 681 | 14 | 1149 | 21 | AA60917 | Phanerocchaete sord | C 754 | 14 | 2.4 | 1428 | 20 | AA708799 | Human homologue of |
| C 682 | 14 | 1159 | 21 | AA60917 | Human secreted pro | C 755 | 14 | 2.4 | 1437 | 18 | AA770130 | Max-interacting pr |
| C 683 | 14 | 1161 | 21 | AA60917 | Arbidopsin thalia | C 756 | 14 | 2.4 | 1437 | 21 | AA705493 | Streptococcus pneu |
| C 684 | 14 | 1164 | 19 | AA735131 | Mouse WRN helicase | C 757 | 14 | 2.4 | 1446 | 21 | AA53032 | Murine beta-1,3-ga |
| C 685 | 14 | 1164 | 21 | AA735131 | Arbidopsin thalia | C 758 | 14 | 2.4 | 1448 | 21 | AA666434 | Human secreted pro |
| C 686 | 14 | 1170 | 21 | AA648952 | Arbidopsin thalia | C 759 | 14 | 2.4 | 1448 | 21 | AA668152 | Eucalyptus grandis |
| C 687 | 14 | 1172 | 21 | AA648952 | Arbidopsin thalia | C 760 | 14 | 2.4 | 1456 | 21 | AA639078 | Human secreted pro |
| C 688 | 14 | 1175 | 21 | AA63905 | Human secreted pro | C 761 | 14 | 2.4 | 1469 | 20 | AA728654 | Nucleotide sequenc |
| C 689 | 14 | 1193 | 21 | AA63905 | Pinus radiata tran | C 762 | 14 | 2.4 | 1469 | 21 | AA440574 | Xenopus sp embryo |
| C 690 | 14 | 1198 | 21 | AA632888 | Arbidopsin thalia | C 763 | 14 | 2.4 | 1478 | 21 | AA642221 | Arbidopsin thalia |
| C 691 | 14 | 1199 | 21 | AA616672 | Human secreted pro | C 764 | 14 | 2.4 | 1488 | 20 | AA739661 | Renal cancer assoc |
| C 692 | 14 | 1200 | 12 | AA611650 | FB-FB-UK fusion co | C 765 | 14 | 2.4 | 1491 | 20 | AA700070 | Aspergillus oryzae |
| C 693 | 14 | 1200 | 20 | AA774138 | Mouse FLAME-2 cDNA | C 766 | 14 | 2.4 | 1491 | 20 | AA782521 | Aspergillus oryzae |
| C 694 | 14 | 1204 | 21 | AA633604 | Arbidopsin thalia | C 767 | 14 | 2.4 | 1491 | 21 | AA700018 | Aspergillus oryzae |
| C 695 | 14 | 1205 | 21 | AA608110 | Human ATP synthase | C 768 | 14 | 2.4 | 1498 | 21 | AA786796 | Human protein kina |
| C 696 | 14 | 1205 | 21 | AA735906 | Human ATP synthase | C 769 | 14 | 2.4 | 1515 | 21 | AA770118 | Plasmodium falci |
| C 697 | 14 | 1209 | 17 | AA739517 | Flea calreticulin | C 770 | 14 | 2.4 | 1518 | 19 | AA734595 | M. vaccae puta hom |
| C 698 | 14 | 1209 | 17 | AA739517 | Flea calreticulin | C 771 | 14 | 2.4 | 1518 | 20 | AA711330 | M. vaccae puta gen |
| C 699 | 14 | 1212 | 21 | AA740247 | Arbidopsin thalia | C 772 | 14 | 2.4 | 1519 | 21 | AA799893 | Human secreted pro |
| C 700 | 14 | 1221 | 21 | AA740247 | Arbidopsin thalia | C 773 | 14 | 2.4 | 1521 | 19 | AA701533 | Human acylcoenzyme |
| C 701 | 14 | 1225 | 21 | AA740247 | Arbidopsin thalia | C 774 | 14 | 2.4 | 1527 | 22 | AA731528 | C. glutamicum phosph |
| C 702 | 14 | 1228 | 22 | AA733098 | Human secreted pro | C 775 | 14 | 2.4 | 1529 | 21 | AA759076 | Sequence from PATK |
| C 703 | 14 | 1234 | 15 | AA704415 | Novel protein kina | C 776 | 14 | 2.4 | 1530 | 21 | AA750275 | Human melancortin |
| C 704 | 14 | 1245 | 22 | AA744651 | Human secreted pro | C 777 | 14 | 2.4 | 1530 | 21 | AA751831 | Human melancortin |
| C 705 | 14 | 1258 | 22 | AA733083 | Human secreted pro | C 778 | 14 | 2.4 | 1530 | 21 | AA751835 | Human melancortin |
| C 706 | 14 | 1261 | 16 | AA780230 | Rat NDF clone 40 D | C 779 | 14 | 2.4 | 1532 | 21 | AA753022 | Human beta-1,3-gal |
| C 707 | 14 | 1268 | 16 | AA738969 | Sequence of delet | C 780 | 14 | 2.4 | 1533 | 21 | AA750047 | Human secreted pro |
| C 708 | 14 | 1268 | 16 | AA701482 | Tobacco promoter T | C 781 | 14 | 2.4 | 1536 | 21 | AA755826 | S. lavendulae Mmcl |
| C 709 | 14 | 1268 | 18 | AA774891 | Human neurogenic d | C 782 | 14 | 2.4 | 1537 | 21 | AA745667 | Arbidopsin thalia |
| C 710 | 14 | 1268 | 19 | AA742932 | DNA encoding human | C 783 | 14 | 2.4 | 1540 | 21 | AA751833 | Human melancortin |
| C 711 | 14 | 1269 | 14 | AA748087 | Sequence of the an | C 784 | 14 | 2.4 | 1540 | 21 | AA751837 | Human melancortin |
| C 712 | 14 | 1269 | 21 | AA742299 | Arbidopsin thalia | C 785 | 14 | 2.4 | 1544 | 21 | AA733610 | Arbidopsin thalia |
| C 713 | 14 | 1270 | 15 | AA744340 | Sequence encoding | C 786 | 14 | 2.4 | 1547 | 21 | AA747981 | Arbidopsin thalia |
| C 714 | 14 | 1272 | 21 | AA733124 | Arbidopsin thalia | C 787 | 14 | 2.4 | 1553 | 21 | AA7297082 | Human secreted pro |
| C 715 | 14 | 1274 | 21 | AA748616 | Arbidopsin thalia | C 788 | 14 | 2.4 | 1586 | 21 | AA790327 | Agrobacterium KNR7 |
| C 716 | 14 | 1275 | 19 | AA705044 | DNA encoding the v | C 789 | 14 | 2.4 | 1589 | 17 | AA739516 | Flea calreticulin |
| C 717 | 14 | 1275 | 19 | AA705044 | DNA encoding the v | C 790 | 14 | 2.4 | 1589 | 17 | AA739518 | Flea calreticulin |
| C 718 | 14 | 1287 | 21 | AA76396 | Human ORFX ORF1951 | C 791 | 14 | 2.4 | 1591 | 20 | AA722715 | Human cytochrome p |
| C 719 | 14 | 1287 | 22 | AA771411 | Corynebacterium g1 | C 792 | 14 | 2.4 | 1593 | 20 | AA701021 | DNA sequence of 5N |
| C 720 | 14 | 1289 | 16 | AA709001 | Bovine herpes viru | C 793 | 14 | 2.4 | 1599 | 21 | AA780531 | Human secreted pro |
| C 721 | 14 | 1290 | 11 | AA703061 | Human T-cell leuk | C 794 | 14 | 2.4 | 1601 | 18 | AA761395 | T72 promoter regio |
| C 722 | 14 | 1305 | 21 | AA716092 | Human prostate can | C 795 | 14 | 2.4 | 1605 | 18 | AA767531 | H. pylori putative |
| C 723 | 14 | 1305 | 21 | AA746454 | Arbidopsin thalia | C 796 | 14 | 2.4 | 1607 | 21 | AA751800 | Arbidopsin thalia |
| C 724 | 14 | 1317 | 18 | AA799937 | Bovine dipeptidyl | C 797 | 14 | 2.4 | 1607 | 21 | AA754052 | Human endothelial |
| C 725 | 14 | 1321 | 21 | AA77590 | Human ORFX ORF3145 | C 798 | 14 | 2.4 | 1608 | 20 | AA760803 | Human secreted pro |
| C 726 | 14 | 1332 | 21 | AA76757 | Nuclear transpor | C 799 | 14 | 2.4 | 1617 | 18 | AA767947 | H. pylori putative |
| C 727 | 14 | 1339 | 21 | AA799968 | Human secreted pro | C 800 | 14 | 2.4 | 1621 | 21 | AA765004 | Membrane-bound pro |
| C 728 | 14 | 1348 | 21 | AA744931 | Arbidopsin thalia | C 801 | 14 | 2.4 | 1621 | 22 | AA744150 | Human Prol134 (UNG |
| C 729 | 14 | 1351 | 21 | AA766800 | Human ORFX ORF2355 | C 802 | 14 | 2.4 | 1633 | 14 | AA751227 | Human MSH-R gene |
| C 730 | 14 | 1353 | 18 | AA784985 | Burkholderia cepa | C 803 | 14 | 2.4 | 1633 | 19 | AA719136 | Human melanocyte s |
| C 731 | 14 | 1371 | 21 | AA768887 | Human ORFX ORF2442 | C 804 | 14 | 2.4 | 1636 | 20 | AA763303 | Human melanocyte s |
| C 732 | 14 | 1376 | 21 | AA739392 | Arabidopsis thalia | C 805 | 14 | 2.4 | 1636 | 21 | AA777774 | Human cancer assoc |
| C 733 | 14 | 1380 | 18 | AA751051 | Human amine recept | C 806 | 14 | 2.4 | 1640 | 21 | AA745188 | Aspergillus sojae |
| C 734 | 14 | 1380 | 20 | AA781288 | Human amine recept | C 807 | 14 | 2.4 | 1650 | 15 | AA774684 | Early ripening Tom |
| C 735 | 14 | 1383 | 19 | AA742055 | Kurthia sp. biotin | C 808 | 14 | 2.4 | 1653 | 22 | AA774435 | Human pro4 nucleot |
| C 736 | 14 | 1385 | 14 | AA738968 | Sequence of delet | C 809 | 14 | 2.4 | 1659 | 22 | AA761037 | P. putida KT2440-a |
| C 737 | 14 | 1385 | 16 | AA701481 | Tobacco promoter T | C 810 | 14 | 2.4 | 1663 | 20 | AA733162 | Potato isomylase |
| C 738 | 14 | 1385 | 22 | AA771410 | Corynebacterium g1 | C 811 | 14 | 2.4 | 1671 | 21 | AA739977 | Human secreted pro |
| C 739 | 14 | 1386 | 18 | AA772332 | Pseudomonas aerugi | C 812 | 14 | 2.4 | 1673 | 22 | AA728835 | Arbidopsin thalia |
| C 740 | 14 | 1387 | 22 | AA726301 | Pseudomonas sp. TYP | C 813 | 14 | 2.4 | 1675 | 21 | AA742587 | Arbidopsin thalia |
| C 741 | 14 | 1391 | 22 | AA793793 | Human CDNA encodin | C 814 | 14 | 2.4 | 1680 | 19 | AA739818 | Scenedesmus D1 pro |

| | | | | | | |
|-------|----|-----|------|----|-----------|---------------------|
| C 815 | 14 | 2.4 | 1689 | 21 | AAZ29898 | Human Rhotekin pro |
| C 816 | 14 | 2.4 | 1689 | 22 | AAF24437 | Human rhotekin cod |
| C 817 | 14 | 2.4 | 1694 | 21 | AAAF2016 | Hydrophobic domain |
| C 818 | 14 | 2.4 | 1695 | 17 | AAAT36138 | Guinea pig L-aspar |
| C 819 | 14 | 2.4 | 1708 | 18 | AAAT89519 | Human cytokine Ler |
| C 820 | 14 | 2.4 | 1722 | 21 | AAZ98117 | Human signal pepti |
| C 821 | 14 | 2.4 | 1726 | 21 | AAZ36420 | HIV encoding a mai |
| C 822 | 14 | 2.4 | 1733 | 15 | AAO73300 | HIV-type virus MNP |
| C 823 | 14 | 2.4 | 1733 | 18 | AAZ23381 | HIV isolate 5180 g |
| C 824 | 14 | 2.4 | 1735 | 20 | AAAT7840 | Human melanoma ass |
| C 825 | 14 | 2.4 | 1740 | 21 | AAZ36914 | DNA encoding an ac |
| C 826 | 14 | 2.4 | 1740 | 22 | AAAC85084 | Atherosclerosis-as |
| C 827 | 14 | 2.4 | 1743 | 15 | AAO71737 | Mutant gamma-glu |
| C 828 | 14 | 2.4 | 1743 | 21 | AAZ86927 | Netrin coding sequ |
| C 829 | 14 | 2.4 | 1746 | 22 | AAZ56603 | H. pylori HPS013 e |
| C 830 | 14 | 2.4 | 1766 | 18 | AAAT91131 | Strawberry fruit c |
| C 831 | 14 | 2.4 | 1770 | 22 | AAAF61373 | A. Challana chloro |
| C 832 | 14 | 2.4 | 1776 | 21 | AAAA49177 | cDNA encoding huma |
| C 833 | 14 | 2.4 | 1779 | 21 | AAZ65347 | Human secreted pro |
| C 834 | 14 | 2.4 | 1780 | 21 | AAAF21665 | Human breast and o |
| C 835 | 14 | 2.4 | 1792 | 22 | AAAT7810 | Secreted protein g |
| C 836 | 14 | 2.4 | 1798 | 18 | AAAT51756 | DXN-1 gene promote |
| C 837 | 14 | 2.4 | 1801 | 21 | AAZ36893 | cDNA encoding an a |
| C 838 | 14 | 2.4 | 1805 | 21 | AAAC41842 | Arabidopsis thalia |
| C 839 | 14 | 2.4 | 1818 | 11 | AAAO3599 | Human liver cytoch |
| C 840 | 14 | 2.4 | 1818 | 21 | AAAT72435 | Human nuclear acid |
| C 841 | 14 | 2.4 | 1820 | 20 | AAV59096 | TFAS protein codi |
| C 842 | 14 | 2.4 | 1830 | 22 | AAAC84343 | Wheat Np2-2 prote |
| C 843 | 14 | 2.4 | 1833 | 19 | AAV46339 | Serine threonine k |
| C 844 | 14 | 2.4 | 1836 | 22 | AAAD02610 | Human angipolelin |
| C 845 | 14 | 2.4 | 1839 | 21 | AAAT5087 | Open reading frame |
| C 846 | 14 | 2.4 | 1841 | 20 | AAZ23022 | Human kd312 polype |
| C 847 | 14 | 2.4 | 1845 | 20 | AAAT19925 | Mammalian cytochro |
| C 848 | 14 | 2.4 | 1845 | 21 | AAZ98400 | Canine beta1-adren |
| C 849 | 14 | 2.4 | 1848 | 21 | AAAT57771 | Soybean putative c |
| C 850 | 14 | 2.4 | 1848 | 21 | AAZ40671 | Human presenilin-2 |
| C 851 | 14 | 2.4 | 1852 | 17 | AAAT11381 | Cytochrome P450 2C |
| C 852 | 14 | 2.4 | 1852 | 19 | AAV44157 | Human cytochrome P |
| C 853 | 14 | 2.4 | 1854 | 17 | AAAT11378 | Cytochrome P450 2C |
| C 854 | 14 | 2.4 | 1854 | 19 | AAV44154 | Human cytochrome P |
| C 855 | 14 | 2.4 | 1856 | 21 | AAAC58635 | Human secreted pro |
| C 856 | 14 | 2.4 | 1859 | 21 | AAO38967 | Sequence of deleti |
| C 857 | 14 | 2.4 | 1859 | 16 | AAAT01480 | Tobacco promoter T |
| C 858 | 14 | 2.4 | 1860 | 18 | AAAT69808 | EPH family ligand |
| C 859 | 14 | 2.4 | 1863 | 18 | AAAT47705 | Human FRX1 gene. |
| C 860 | 14 | 2.4 | 1866 | 21 | AAZ65006 | Membrane-bound pro |
| C 861 | 14 | 2.4 | 1866 | 21 | AAZ65267 | Human secreted pro |
| C 862 | 14 | 2.4 | 1866 | 22 | AAAF4152 | Human EST DNA5725 |
| C 863 | 14 | 2.4 | 1872 | 21 | AAAT12401 | cDNA encoding a hu |
| C 864 | 14 | 2.4 | 1872 | 21 | AAZ51250 | Human RNA-associat |
| C 865 | 14 | 2.4 | 1877 | 18 | AAV06354 | AL-2-long (AL-21) |
| C 866 | 14 | 2.4 | 1877 | 18 | AAAT79705 | Human secreted pro |
| C 867 | 14 | 2.4 | 1882 | 21 | AAAC46935 | Arabidopsis thalia |
| C 868 | 14 | 2.4 | 1883 | 19 | AAV52044 | Helicobacter polyp |
| C 869 | 14 | 2.4 | 1894 | 14 | AAO61796 | Human NR5F CDNA. |
| C 870 | 14 | 2.4 | 1894 | 16 | AAO80224 | Rat NDF clone 44 D |
| C 871 | 14 | 2.4 | 1895 | 21 | AAAB8842 | Human acyl CoA:cho |
| C 872 | 14 | 2.4 | 1895 | 21 | AAZ45383 | DNA encoding a pro |
| C 873 | 14 | 2.4 | 1907 | 20 | AAZ10848 | Short-chain tumour |
| C 874 | 14 | 2.4 | 1917 | 21 | AAAT37848 | Human obesity prot |
| C 875 | 14 | 2.4 | 1920 | 21 | AAAT36140 | Rice glutamy1-LRNA |
| C 876 | 14 | 2.4 | 1928 | 17 | AAAT36140 | Guinea pig L-aspar |
| C 877 | 14 | 2.4 | 1928 | 19 | AAV15790 | CDNA for wild type |
| C 878 | 14 | 2.4 | 1933 | 14 | AAO38966 | Sequence of TobRB7 |
| C 879 | 14 | 2.4 | 1933 | 16 | AAAT01479 | Tobacco promoter T |
| C 880 | 14 | 2.4 | 1935 | 21 | AAAT79715 | Eucahyptus grandis |
| C 881 | 14 | 2.4 | 1942 | 21 | AAAC58638 | Human PRO1434 prot |
| C 882 | 14 | 2.4 | 1942 | 21 | AAAT51262 | Human DNA encoding |
| C 883 | 14 | 2.4 | 1944 | 21 | AAAC48428 | Arabidopsis thalia |
| C 884 | 14 | 2.4 | 1950 | 21 | AAAC40701 | Arabidopsis thalia |
| C 885 | 14 | 2.4 | 1958 | 11 | AAO05639 | Clone hps11 encodi |
| C 886 | 14 | 2.4 | 1960 | 11 | AAAT76900 | Human ORFX ORF2455 |
| C 887 | 14 | 2.4 | 1973 | 10 | AAAT91045 | Gene encoding bov1 |
| C 888 | 14 | 2.4 | 1973 | 21 | AAZ29899 | Human Rhotekin var |
| C 889 | 14 | 2.4 | 1973 | 22 | AAAF24438 | Human rhotekin EST |
| C 890 | 14 | 2.4 | 1976 | 21 | AAAT76169 | Human ACAT Related |
| C 891 | 14 | 2.4 | 1978 | 21 | AAZ52432 | HTRM clone 693452 |
| C 892 | 14 | 2.4 | 1979 | 21 | AAZ56766 | Human transmembran |
| C 893 | 14 | 2.4 | 1983 | 21 | AAZ40670 | Human presenilin-2 |
| C 894 | 14 | 2.4 | 1989 | 16 | AAZ26147 | Rat pan-s/tk recep |
| C 895 | 14 | 2.4 | 1996 | 20 | AAO92805 | Alpha-Amy6-C gene. |
| C 896 | 14 | 2.4 | 1996 | 21 | AAAT7946 | Rice alpha-Amy6-C |
| C 897 | 14 | 2.4 | 2000 | 13 | AAO21400 | Vibrolysin gene se |
| C 898 | 14 | 2.4 | 2000 | 20 | AAV84700 | Neutral protease v |
| C 899 | 14 | 2.4 | 2002 | 21 | AAZ40669 | Human presenilin-2 |
| C 900 | 14 | 2.4 | 2019 | 21 | AAAT95815 | Coron MFPI EST. ze |
| C 901 | 14 | 2.4 | 2022 | 18 | AAAT51755 | DXN-1 gene associa |
| C 902 | 14 | 2.4 | 2024 | 14 | AAZ06229 | Human secreted pro |
| C 903 | 14 | 2.4 | 2025 | 21 | AAAC51726 | zee mays DNA fragm |
| C 904 | 14 | 2.4 | 2027 | 19 | AAV23979 | Human CD33-like pr |
| C 905 | 14 | 2.4 | 2030 | 12 | AAO11651 | PB-PB-SK fusion co |
| C 906 | 14 | 2.4 | 2040 | 21 | AAAT52773 | Soybean putative c |
| C 907 | 14 | 2.4 | 2040 | 22 | AAAF74278 | Rice calcium depen |
| C 908 | 14 | 2.4 | 2042 | 18 | AAO46127 | Caulliflower consti |
| C 909 | 14 | 2.4 | 2042 | 18 | AAAT63693 | Brassica heat choc |
| C 910 | 14 | 2.4 | 2053 | 21 | AAAT27050 | Human cell surface |
| C 911 | 14 | 2.4 | 2055 | 21 | AAAC42837 | Arabidopsis thalia |
| C 912 | 14 | 2.4 | 2079 | 21 | AAAT26360 | Human secreted pro |
| C 913 | 14 | 2.4 | 2110 | 20 | AAZ40535 | Full length human |
| C 914 | 14 | 2.4 | 2115 | 19 | AAV42799 | Rat CMV IE2 plasm |
| C 915 | 14 | 2.4 | 2120 | 22 | AAAC84345 | Wheat Np2-2 cDNA |
| C 916 | 14 | 2.4 | 2124 | 9 | AAAT82101 | Complementary DNA |
| C 917 | 14 | 2.4 | 2126 | 21 | AAAT10312 | European corn bore |
| C 918 | 14 | 2.4 | 2133 | 18 | AAAT96754 | S. tuberosum debra |
| C 919 | 14 | 2.4 | 2144 | 21 | AAZ40668 | Human presenilin-2 |
| C 920 | 14 | 2.4 | 2151 | 20 | AAZ12020 | Neisseria gonorrhoe |
| C 921 | 14 | 2.4 | 2159 | 15 | AAO67223 | Mouse p53Nuc. Mus |
| C 922 | 14 | 2.4 | 2170 | 21 | AAAT99328 | Human Ras correlat |
| C 923 | 14 | 2.4 | 2186 | 16 | AAO51344 | Rat NNSF CDNA. Ra |
| C 924 | 14 | 2.4 | 2186 | 16 | AAO80223 | Rat NDF DNA. Rat |
| C 925 | 14 | 2.4 | 2192 | 22 | AAAF32727 | Human secreted pro |
| C 926 | 14 | 2.4 | 2192 | 18 | AAAT77838 | Human melanoma ass |
| C 927 | 14 | 2.4 | 2208 | 12 | AAO11140 | Cellular Receptor |
| C 928 | 14 | 2.4 | 2208 | 20 | AAZ52917 | Human prostate tum |
| C 929 | 14 | 2.4 | 2210 | 21 | AAAT57827 | Human ORFX ORF1382 |
| C 930 | 14 | 2.4 | 2217 | 17 | AAAC44902 | Arabidopsis thalia |
| C 931 | 14 | 2.4 | 2229 | 21 | AAAT40031 | Human presenilin-2 |
| C 932 | 14 | 2.4 | 2229 | 19 | AAV04669 | Human presenilin-2 |
| C 933 | 14 | 2.4 | 2236 | 18 | AAAT51253 | Human AdA protein |
| C 934 | 14 | 2.4 | 2236 | 19 | AAAT75762 | Human AdA protein |
| C 935 | 14 | 2.4 | 2243 | 22 | AAAC89709 | Human presenilin I |
| C 936 | 14 | 2.4 | 2251 | 21 | AAAC64971 | Human oligodendroc |
| C 937 | 14 | 2.4 | 2271 | 20 | AAAT19485 | C. elegans sarco/e |
| C 938 | 14 | 2.4 | 2276 | 18 | AAAT87426 | Human secreted pro |
| C 939 | 14 | 2.4 | 2295 | 21 | AAAT80610 | Full AD4/AD31p seq |
| C 940 | 14 | 2.4 | 2301 | 21 | AAAT90221 | Human CD33-like se |
| C 941 | 14 | 2.4 | 2305 | 17 | AAAT16258 | Human pancreatic c |
| C 942 | 14 | 2.4 | 2305 | 20 | AAV63338 | Rice cultivar IR58 |
| C 943 | 14 | 2.4 | 2315 | 12 | AAO12092 | Nucleotide sequenc |
| C 944 | 14 | 2.4 | 2335 | 13 | AAO31911 | Pregnancy-specific |
| C 945 | 14 | 2.4 | 2335 | 13 | AAO31911 | Xenopus activin re |
| C 946 | 14 | 2.4 | 2335 | 22 | AAAX22684 | Xenopus activin re |
| C 947 | 14 | 2.4 | 2336 | 20 | AAAC85298 | Xenopus activin re |
| C 948 | 14 | 2.4 | 2336 | 16 | AAAT03824 | Bovine herpesviral |
| C 949 | 14 | 2.4 | 2336 | 19 | AAV42798 | Rat CMV IE1 plasm |
| C 950 | 14 | 2.4 | 2338 | 16 | AAAX90030 | Rat connective tis |
| C 951 | 14 | 2.4 | 2356 | 20 | AAO80229 | Rat NDF clone 38 D |
| C 952 | 14 | 2.4 | 2367 | 18 | AAAT84986 | Burkholderia cepac |
| C 953 | 14 | 2.4 | 2377 | 10 | AAAT90726 | DNA sequence of ex |
| C 954 | 14 | 2.4 | 2377 | 15 | AAO68538 | Extracellular neut |
| C 955 | 14 | 2.4 | 2380 | 18 | AAV06355 | AL-2-short (AL-2s) |
| C 956 | 14 | 2.4 | 2397 | 21 | AAAO8835 | Human APC-2 gene. |
| C 957 | 14 | 2.4 | 2403 | 21 | AAAT37095 | Human PRO1472 (UNO |
| C 958 | 14 | 2.4 | 2403 | 22 | AAAF92109 | Human PRO1472 CDNA |
| C 959 | 14 | 2.4 | 2403 | 22 | AAAF34392 | Probe #40 used in |
| C 960 | 14 | 2.4 | 2408 | 16 | AAO81334 | (1-3)-beta-D-gluca |
| C 961 | 14 | 2.4 | 2425 | 22 | AAAF45073 | Human secreted pro |

| DT | 26-NOV-1996 | (first entry) |
|----|---|---|
| XX | Trypanosoma cruzi epimastigotic Prc100t antigen gene. | |
| XX | Antigen: Trypanosoma cruzi; epimastigote; serum; Chagas disease; probe; | |
| KW | Primer; PCR; polymerase chain reaction; amplification; antibody; ds. | |
| XX | Trypanosoma cruzi. | |
| OS | | |
| XX | Key | Location/Qualifiers |
| XX | CDS | 266..3013 |
| XX | | /*tag= a |
| XX | | /product= Prc100t epimastigotic antigen |
| XX | FR2723589-A1. | |
| XX | 16-FEB-1996. | |
| XX | 12-AUG-1994; | 94FR-0010132. |
| XX | 12-AUG-1994; | 94FR-0010132. |
| XX | (INMR) BIO MERIEUX. | |
| XX | Jolivet M, Lesenechal M, Paranhos-Baccala G; | |
| XX | WPI; 1996-190287/20. | |
| XX | P-PSDB; AAR91615. | |
| XX | New nucleic acid encoding Trypanosoma cruzi epimastigotic antigen - | |
| XX | useful for diagnosis, monitoring and therapy of Chagas disease | |
| XX | Claim 1; Page 24-26; 55pp; French. | |
| XX | This is the nucleotide sequence encoding a novel isolated antigenic | |
| XX | protein from Trypanosoma cruzi epimastigotes, designated Prc100t. | |
| XX | The clone Tc50 was isolated from a T. cruzi genomic expression library in | |
| XX | Lambda gt11, using a mixture of sera from patients with Chagas disease. | |
| XX | Clone Tc50 contained an 594 bp insert corresp. to nucleotides 1232-1825 | |
| XX | of this sequence. The Tc50 sequence was subsequently used to probe a | |
| XX | Southern blot of restriction enzyme digested T. cruzi DNA and also screen | |
| XX | a lambda gt10 library to isolate a 1041 bp EcoRI fragment corresp. to | |
| XX | nucleotides 1403-2443 of Prc100t. Primers (AAT7311-5) were synthesised | |
| XX | based on the sequences of the 594 and 1041 bp fragments and used to | |
| XX | amplify the Prc100t clone as 3 fragments from cDNA derived from mRNA | |
| XX | purified from T. cruzi epimastigotes. The protein or antibodies raised | |
| XX | against it can be used in the detection and monitoring of T. cruzi | |
| XX | infection i.e. Chagas disease. | |
| XX | Sequence 3402 BP; 889 A; 818 C; 958 G; 737 T; 0 other; | |
| XX | Query Match | 82.8%; Score 492; DB 17; Length 3402; |
| XX | Best Local Similarity | 99.7%; Pred. No. 3e-243; |
| XX | Matches 592; Conservative | 0; Mismatches 2; Indels 0; Gaps 0 |
| QY | 1 caggatcagcgtaagcgcttctgtcttcaatcgttaacgcccagcgtatgcgtctcgtgct 60 | |
| Db | 1232 caggtacgcgctaaagcgtttgtcttcaatcgttaacgcccagcgtatgcgtctcgtgct 1291 | |
| QY | 61 gacatgtcgtatcgtatgaacatcttccatctcgcgtctcctccgcagggaaacacagcag 120 | |
| Db | 1292 gacatgtcgtatcgtatgaacatcttccatctcgcgtctcctccgcagggaaacacagcag 1351 | |
| QY | 121 ccaggccaaaaaacctcggtagtggcgcgcgcgggaacccggggtgtgtgtcctccggcgact 180 | |
| Db | 1352 ccaggccaaaaaacctcggtagtggcgcgcgcgggaacccggggtgtgtgtcctccggcgact 1411 | |
| QY | 181 gacgcgcgagtagacgcatcatccataacgaactcttgcgcgcgtcgcgcgtcccttgatca 240 | |
| Db | 1412 gacgcgcgagtagacgcatcatccataacgaactcttgcgcgcgtcgcgcgtcccttgatca 1471 | |
| QY | 241 cccctgtttcagcgccagcaagcgcgcgcgtcctctgcgcgcgcgcgcgtatgcgttgag 300 | |

| | | | |
|----|------|--|------|
| Db | 1472 | ccccctgcttcacgagccagacggaagcagccgccccctccgccccgagcgacacgccccctgag | 1533 |
| Oy | 301 | ccgcacgctgaggagccaagatcatctgtcaactctagtgaaatcagctgggagataatgtacac | 360 |
| Db | 1532 | ccgcacgctggggagcaagatcatctgtcaactctagtgaaatcagctgggagataatgtacac | 1591 |
| Oy | 361 | caaaaggagctctgctcgaacgactctgagagcgccggccaagaaagagctctcagcggttgacgtcc | 420 |
| Db | 1592 | caaaaggagctctgctcgaacgactctgagagcgccggccaagaaagagctctcagcggttgacgtcc | 1651 |
| Oy | 421 | acgactacgcgcccgccgacgcaaacagatgcatacgggacacaaatgagccgagactctgtacagct | 480 |
| Db | 1652 | acgactacgcgcccgccgacgcaaacagatgcatacgggacacaaatgagccgagactctgtacagct | 1711 |
| Oy | 481 | ggatgtgtgagcagctaaatagtgtgtgacagcgcgccctgctctccaaagccgacggcgaa | 540 |
| Db | 1712 | ggatgtgtgagcagctaaatagtgtgtgacagcgcgccctgctctccaaagccgacggcgaa | 1773 |
| Oy | 541 | ccaacagagagagaaagaaagcgctcccgcgagatgtgaaagcagctccctgttcgcgata | 594 |
| Db | 1772 | ccaacagagagagaaagaaagcgctcccgcgagatgtgaaagcagctccctgttcgcgata | 1825 |

| | | |
|----|----------|--|
| XX | RESULT | 3 |
| XX | AT27313 | |
| XX | ID | AA2727313 standard; cDNA, 22 BP. |
| XX | AC | |
| XX | AA27313: | |
| XX | DT | 26-NOV-1996 (first entry) |
| XX | DE | T.cruzi epimastigotic PTC100t antigen primer corresp. to bases 1266-87. |
| XX | XX | |
| XX | XX | Antigen; Trypanosoma cruzi; epimastigote; serum; Chagas disease; probe; |
| XX | XX | Primer; PCR; polymerase chain reaction; amplification; antibody; ss. |
| XX | OS | Synthetic. |
| XX | XX | |
| XX | XX | FR2723589-A1. |
| XX | XX | |
| XX | XX | 16-FEB-1996. |
| XX | PD | |
| XX | XX | |
| XX | PF | 12-AUG-1994; 94FR-0010132. |
| XX | XX | |
| XX | PR | 12-AUG-1994; 94FR-0010132. |
| XX | XX | |
| XX | PA | (INMR) BIO MERIEUX. |
| XX | XX | |
| XX | PI | Jolivet M, Lesenechal M, Paranhos-Baccala G; |
| XX | XX | |
| XX | XX | WPI: 1996-190287/20. |
| XX | DR | |
| XX | PT | New nucleic acid encoding Trypanosoma cruzi epimastigotic antigen - |
| XX | PT | useful for diagnosis, monitoring and therapy of Chagas disease |
| XX | XX | |
| XX | PS | Claim 29; Page 37; 55pp; French. |
| XX | XX | |
| XX | CC | The primers AA27311-5 were used to PCR amplify the sequence encoding a |
| XX | CC | novel isolated antigenic protein from Trypanosoma cruzi epimastigotes, |
| XX | CC | designated PTC100t (AA27310). The primers, derived from the sequences |
| XX | CC | of a 594 and 1041 bp fragment of PTC100t, amplified the gene as 3 |
| XX | CC | fragments. This primer corresponds to nucleotides 1266-87 of the PTC100t |
| XX | CC | sequence, derived from the 594 bp fragment. The 594 bp fragment was |
| XX | CC | isolated from a T.cruzi genomic expression library in lambda g11, using |
| XX | CC | a mixture of sera from patients with Chagas disease. It corresponds to |
| XX | CC | nucleotides 1237-1825 of PTC100t. The 1041 bp fragment was isolated from |
| XX | CC | a lambda g110 library using the 594 bp fragment as a probe. The protein, |
| XX | CC | or antibodies raised against it, can be used in the detection and |
| XX | CC | monitoring of T.cruzi infection i.e. Chagas disease. |
| XX | XX | |
| XX | Sequence | 22 BP; 3 A; 8 G; 7 G; 4 T; 0 other; |

Query Match 3.7%; Score 22; DB 17; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 35 cagcgacggtagctgctct 56
 ||||||||||||||||||
 DB 1 cagcgacggtagctgctct 22

RESULT 4

AAV30199 standard; DNA: 22 BP.

AC AAV30199;

DE 27-AUG-1999 (first entry)

XX PCR primer for T. cruzi PTC40 coding sequence.

XX PTC40; Tc40; infection; diagnosis; immune complex; antigenic determinant;
 KM therapy; antibody; PCR primer; ss.

XX Synthetic.

OS Trypanosoma cruzi.

XX WO99239867-A1.

PD 17-JUN-1999.

PF 10-DEC-1998; 98WO-IB01987.

PR 10-DEC-1997; 97US-0988242.

PA (INMR) BIO MERIEUX.

PI Jolivet M, Lesenechal M, Mandrand B, Paranhos-Baccala G;

DR WPI; 1999-394978/33.

PT New Trypanosoma cruzi antigen

PS Disclosure; Page 21; 65pp; English.

CC This sequence is a PCR primer for DNA encoding the Trypanosoma cruzi

CC PTC40 protein of the invention. The PTC40 antigenic determinant is

CC useful as a reagent for detection and/or monitoring of Trypanosoma cruzi

CC infection from samples including blood serum or plasma, urine, saliva, or

CC tears, by contacting with the sample and detecting an immune complex. The

CC PTC40 antigenic determinant, the vector, expression cassette, cell or

CC antibody are useful for treatment or prevention (vaccine) of a

CC Trypanosoma cruzi infection in a man or animal. Current Trypanosoma cruzi

CC antigens are obtained from protein fractions of the noninfectious stage

CC of the parasite, and these do not allow sufficient production of antigens

CC for use in reliable serological diagnostic tests. The strain to strain

CC polymorphism reduces reliability of the tests.

Sequence 22 BP; 3 A; 8 C; 7 G; 4 T; 0 other;

Query Match 3.7%; Score 22; DB 20; Length 22;

Best Local Similarity 100.0%; Pred. No. 0.12; Mismatches 0; Indels 0; Gaps 0;

OY 35 cagcgacggtagctgctct 56
 ||||||||||||||||||
 DB 1 cagcgacggtagctgctct 22

RESULT 5

AAV30199/c

ID AAV30199 standard; DNA: 11878 BP.

AC AAV30199;
 XX 14-SEP-1998 (first entry)

DE Protein kinase catalytic subunit gene.

XX Severe combined immunodeficiency disease; SCID; horse; diagnosis;
 KM DNA-dependent protein kinase; ds.

OS Equus caballus.

PN WO9821367-A1.

PD 22-MAY-1998.

PF 14-NOV-1997; 97WO-US21066.

PR 15-NOV-1996; 96US-0031261.

PA (TEXA) UNIV TEXAS SYSTEM.

PI Weeks K;

DR WPI; 1998-297967/26.

PT DNA-dependent protein kinase catalytic subunit - useful for
 determining equine severe combined immunodeficiency alleles

PS Disclosure; Page 70-78; 98pp; English.

CC This isolated DNA molecule encodes an equine DNA-dependent protein

CC kinase catalytic subunit (DNA-PKcs). A claimed method of

CC identifying an Arabian horse that is a carrier of equine severe

CC combined immunodeficiency (SCID) comprises determining whether the

CC horse has a mutation in a SCID determinant region of the DNA-PKcs

CC gene (see also AAV30196 and AAV30197). Sequence analysis of DNA-PKcs

CC genes from normal and SCID equine fibroblasts shows that a 5 bp

CC deletion is present in SCID foals at a site that corresponds to

CC nucleotide 9454 of the 12,381 nucleotide coding sequence of the

CC human transcript. This results in premature termination of the

CC DNA-PKcs at amino acid 3160 (see AAV56642). Oligonucleotide probes

CC (see AAV30194 and AAV30195) are provided that precisely span the SCID

CC determinant region of the DNA-PKcs gene, and which are diagnostic

CC for the normal and SCID alleles, respectively. Methods are also

CC provided for identifying for differentiating SCID homozygotes,
 CC heterozygotes and normal horses.

Sequence 11878 BP; 3586 A; 2323 C; 2668 G; 3301 T; 0 other;

Query Match 3.2%; Score 19; DB 19; Length 11878;

Best Local Similarity 100.0%; Pred. No. 3.4; Mismatches 0; Indels 0; Gaps 0;

OY 211 actctgcgctgctcat 229
 ||||||||||||||||||
 DB 6446 ACTCTGCCCTGCTCAT 6428

RESULT 6

AAV30198/c

ID AAV30198 standard; DNA: 11883 BP.

AC AAV30198;

DE 14-SEP-1998 (first entry)

XX Protein kinase catalytic subunit gene.

XX Severe combined immunodeficiency disease; SCID; horse; diagnosis;
 KM DNA-dependent protein kinase; ds.

OS Equus caballus.

```

XX  MO9821367-A1.
PN
XX
XX  22-MAY-1998.
PD
XX
XX  14-NOV-1997; 97WO-US21066.
PF
XX  15-NOV-1996; 96US-0031261.
PR
XX  (TEXA ) UNIV TEXAS SYSTEM.
PA
XX
XX  Weeks K:
PI
XX
XX  WPI: 1998-297967/26.
DR
XX  P-PSDB: AAW56642.
PT
XX  DNA-dependent protein kinase catalytic subunit - useful for
XX  determining equine severe combined immunodeficiency alleles
PS
XX  Claim 1; Page 39-44; 98pp; English.
XX
XX  This isolated DNA molecule encodes a DNA-dependent protein kinase
XX  catalytic subunit (DNA-PKcs, see AAW56642) found in Arabian horses.
XX  A claimed method of identifying an Arabian horse that is a carrier
XX  of equine severe combined immunodeficiency (SCID) comprises
XX  determining whether the horse has a mutation in a SCID determinant
XX  region of the DNA-PKcs gene (see also AAV30196 and AAV30197). Sequence
XX  analysis of DNA-PKcs genes from normal and SCID equine fibroblasts
XX  shows that a 5 bp deletion is present in SCID foals at a site that
XX  corresponds to nucleotide 9454 of the 12,381 nucleotide coding
XX  sequence of the human transcript. This results in premature
XX  termination of the DNA-PKcs at amino acid 3160. Oligonucleotide
XX  probes (see AAV30194 and AAV30195) are provided that precisely span the
XX  SCID determinant region of the DNA-PKcs gene, and which are
XX  diagnostic for the normal and SCID alleles, respectively. Methods
XX  are also provided for identifying for differentiating SCID
XX  homozygotes, heterozygotes and normal horses. A claimed plasmid
XX  contains the DNA-PKcs DNA and regulatory elements necessary for
XX  expression of the DNA in a recombinant cell.
SQ
XX
XX  Sequence 11883 BP; 3588 A; 2324 C; 2669 G; 3302 T; 0 other;

Query Match      3.2%; Score 19; DB 19; Length 11883;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  211 actctgcgcgtcgtcat 229
    |||||||
DB  6449 ACTCTGCCGCTGTCGAT 6431

RESULT 7
AAT27312/c
ID  AAT27312 standard; cDNA; 18 BP.
XX
XX  AAT27312;
AC
XX
XX  26-NOV-1996 (first entry)
DT
XX
XX  T. cruzi epimastigotic PTC100t antigen primer corresp. to bases 1442-59.
DE
XX
XX  Antigen: Trypanosoma cruzi; epimastigote; serum; Chagas disease; probe;
KW  primer; PCR; polymerase chain reaction; amplification; antibody; ss.
XX
XX  Synthetic.
OS
XX  FR2723589-A1.
XX
XX  16-FEB-1996.
XX
XX  12-AUG-1994; 94FR-0010132.
XX

```

```

PR  12-AUG-1994; 94FR-0010132.
XX
XX  (INMR ) BIO MERIEUX.
PA
XX
XX  Jolivet M, Lesenechal M, Paranhos-Baccala G;
PI
XX
XX  WPI: 1996-190287/20.
DR
XX
XX  New nucleic acid encoding Trypanosoma cruzi epimastigotic antigen -
PT  useful for diagnosis, monitoring and therapy of Chagas disease
XX
XX  Claim 29; Page 36; 55pp; French.
PS
XX
XX  The primers AAT27311-5 were used to PCR amplify the sequence encoding a
XX  novel isolated antigenic protein from Trypanosoma cruzi epimastigotes,
XX  designated PTC100t (AAT27310). The primers, derived from the sequences
XX  of a 594 and 1041 bp fragment of PTC100t, amplified the gene as 3
XX  fragments. This primer corresponds to nucleotides 1442-59 of the PT100t
XX  sequence, derived from the 594 bp fragment. The 594 bp fragment was
XX  isolated from a T. cruzi genomic expression library in lambda gtl1, using
XX  a mixture of sera from patients with Chagas disease. It corresponds to
XX  nucleotide 1232-1825 of PTC100t. The 1041 bp fragment was isolated from
XX  a lambda gt10 library using the 594 bp fragment as a probe. The protein,
XX  or antibodies raised against it, can be used in the detection and
XX  monitoring of T. cruzi infection i.e. Chagas disease.
SQ
XX
XX  Sequence 18 BP; 5 A; 4 C; 7 G; 2 T; 0 other;

Query Match      3.0%; Score 18; DB 17; Length 18;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  211 actctgcgcgtcgtcat 228
    |||||||
DB  18 ACTCTGCCGCTGTCGCA 1

RESULT 8
AAX84094/c
ID  AAX84094 standard; DNA; 18 BP.
XX
XX  AAX84094;
AC
XX
XX  27-AUG-1999 (first entry)
DT
XX
XX  PCR primer for T. cruzi PTC40 coding sequence.
DE
XX
XX  PTC40; Tcd0; infection; diagnosis; immune complex; antigenic determinant;
KW  therapy; antibody; PCR primer; ss.
XX
XX  Synthetic.
OS
XX  Trypanosoma cruzi.
XX
XX  WO9929867-A1.
XX
XX  17-JUN-1999.
PD
XX
XX  10-DEC-1998; 98WO-IB01987.
XX
XX  10-DEC-1997; 97US-0988242.
XX
XX  (INMR ) BIO MERIEUX.
PA
XX
XX  Jolivet M, Lesenechal M, Mandrand B, Paranhos-Baccala G;
PI
XX
XX  WPI: 1999-394978/33.
XX
XX  New Trypanosoma cruzi antigen
PT
XX
XX  Disclosure: Page 21; 65pp; English.
PS
XX
XX  This sequence is a PCR primer for DNA encoding the Trypanosoma cruzi
XX

```

CC Ptic40 protein of the invention. The Ptic40 antigenic determinant is
 CC useful as a reagent for detection and/or monitoring of *Trypanosoma cruzi*
 CC infection from samples including blood serum or plasma, urine, saliva, or
 CC tears. By contacting with the sample and detecting an immune complex. The
 CC Ptic40 antigenic determinant, the vector, expression cassette, cell or
 CC antibody are useful for treatment or prevention (vaccine) of a
 CC *Trypanosoma cruzi* infection in a man or animal. Current *Trypanosoma cruzi*
 CC antigens are obtained from protein fractions of the noninfectious stage
 CC of the parasite, and these do not allow sufficient production of antigens
 CC for use in reliable serological diagnostic tests. The strain to strain
 CC polymorphism reduces reliability of the tests.
 CC
 CC Sequence 18 BP; 5 A; 4 C; 7 G; 2 T; 0 other;

Query Match 3.0%; Score 18; DB 20; Length 18;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 actctgcgcgcctctgca 228
 ||||||||||||||||
 Db 18 ACTTCTGCCGCTGCTGCA 1

RESULT 9

AAFO9309 standard; cDNA; 454 BP.

AAFO9309;

13-MAR-2001 (first entry)

Fusarium venenatum EST SEQ ID NO:1832.

Multiple gene expression; filamentous fungal cell; EST;

expressed sequence tag; *Fusarium venenatum*; *Aspergillus niger*;

Aspergillus oryzae; *Trichoderma reesei*; identification; recombination;

culture condition; environmental stress; spore morphogenesis;

metabolic pathway engineering; catabolic pathway engineering; ss.

Fusarium venenatum.

WO200056762-A2.

28-SEP-2000.

22-MAR-2000; 2000WO-US07781.

22-MAR-1999; 99US-0273623.

(NOVO) NOVO NORDISK BIOTECH INC.

(NOVO) NOVO NORDISK AS.

Berka RM, Rey MM, Shuster JR, Kaupinen S, Clausen IG, Olsen PB;

WPI; 2000-594572/56.

Monitoring differential expression of genes in filamentous fungal cells

uses fluorescence-labeled nucleic acids isolated from the cells and a

substrate of expressed sequence tags -

Claim 86; Page 1066; 3161pp; English.

The present invention describes a method for monitoring differential
 CC expression of genes in a first filamentous fungal (FF) cell relative to
 CC expression of the same genes in one or more second filamentous fungal
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from
 CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
 CC are used in the methods for monitoring differential expression of genes
 CC in a first filamentous fungal (FF) cell relative to expression of the
 CC same genes in one or more second filamentous fungal cells. Monitoring
 CC the global expression of genes from FF cells allows the production
 CC potential of the microorganisms to be improved. New genes may be

CC discovered, possible functions of unknown open reading frames can be
 CC identified and gene copy number variation and stability can be
 CC monitored. The expression of genes can be used to study how FF cells
 CC adapt to changes in culture conditions, environmental stress, spore
 CC morphogenesis, recombination, metabolic or catabolic pathway
 CC engineering. Using ESTs provides several advantages over genomic or
 CC random cDNA clones including elimination of redundancy as one spot on an
 CC array equals one gene or open reading frame, and organization of the
 CC microarrays based on function of the gene products to facilitate
 CC analysis of the results. AAFO7478 to AAFL1247 represents ESTs from
 CC *Fusarium venenatum*; AAFL1248 to AAFL1853 represents ESTs from *Aspergillus*
 CC *niger*; AAFL1854 to AAFL4878 represents ESTs from *Aspergillus oryzae*; and
 CC AAFL4879 to AAFL5337 represents ESTs from *Trichoderma reesei*, which are
 CC all specifically claimed in the present invention.

Sequence 454 BP; 141 A; 116 C; 89 G; 108 T; 0 other;

Query Match 3.0%; Score 18; DB 21; Length 454;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 220 gctgcgtcaccctctgca 237
 ||||||||||||||||
 Db 205 gctgcgtcaccctctgca 222

RESULT 10

AAFO4783 standard; cDNA; 479 BP.

AAFO4783;

06-OCT-2000 (first entry)

Human secreted protein 5' EST, SEQ ID NO: 8858.

Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

gene therapy; chromosome mapping; ss.

Homo sapiens.

EP1033401-A2.

06-SEP-2000.

21-FEB-2000; 2000EP-0200610.

26-FEB-1999; 99US-0122487.

(GEST) GENSET.

Dumas Mline Edwards J, Duclert A, Giordano J;

WPI; 2000-500381/45.

New nucleic acid that is a 5' expressed sequence tag (5' EST) for

obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for

diagnostic, forensic, gene therapy and chromosome mapping procedures -

Claim 1; SEQ ID 8858; 71pp + CD-ROM; English.

The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used

CC In diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.

SO Sequence 479 BP; 109 A; 109 C; 99 G; 160 T; 2 other;

Query Match 3.0%; Score 18; DB 21; Length 479;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 492 agcataagtggtgccag 509
|||||
DB 64 agcataagtggtgccag 81

RESULT 11
AA39943
ID AAX39943 standard; DNA; 866 BP.

AC AAX39943;

DT 02-JUL-1999 (first entry)

DE Gastric cancer associated gene.

KW Cancer associated antigen; diagnosis; research; treatment; human;
KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW prostate cancer; ss.

OS Homo sapiens.

PN WO9904265-A2.

PD 28-JAN-1999.

PE 15-JUL-1998; 98WO-US14679.

PR 22-JUN-1998; 98US-0102322.

PR 17-JUL-1997; 97US-0896164.

PR 10-OCT-1997; 97US-0061599.

PR 10-OCT-1997; 97US-0061765.

PR 11-OCT-1997; 97US-0948705.

PR 11-OCT-1997; 97GB-0021697.

RA (LUDM-) LUDMIG INST CANCER RES.

PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;

PI Pfeundschnuh M, Sahin U, Scanlan MJ, Stockert E;

PI Tureci O;

PS Claim 67; Page 607; 787pp; English.

CC The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule; and (b) determining the
CC interaction between the agent and the NAM or the expression product as a
CC determination of the disorder. The products and methods can be used in
CC the diagnosis, monitoring, research, or treatment of conditions
CC characterised by the expression of various cancer associated antigens.
CC The invention provides nucleic acid sequences and encoded polypeptides
CC which are cancer associated antigen precursors expressed in human breast
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC lung cancer.

XX
SO Sequence 866 BP; 222 A; 235 C; 197 G; 206 T; 6 other;

Query Match 3.0%; Score 18; DB 20; Length 866;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 230 cccctgcatcccccctg 247
|||||
DB 37 cccctgcatcccccctg 54

RESULT 12
AAC59435/C
ID AAC59435 standard; CDNA; 1302 BP.

AC AAC59435;

DT 02-FEB-2001 (first entry)

DE Human secreted protein CDNA #44.

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antifungal;
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiact; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; ss.

OS Homo sapiens.

PN WO200056765-A1.

PD 28-SEP-2000.

PE 16-MAR-2000; 2000WO-US06823.

PR 19-MAR-1999; 99US-0125364.

PR 08-DEC-1999; 99US-0169623.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulis G;

PI WPT: 2000-602215/57.

PI P-PSDB: AAB34002.

PS Claim 1; Page 355; 410pp; English.

CC The invention relates to the isolation of genes AAC59392-C59439 encoding
CC 48 human secreted proteins AAB33963-B34006. The genes can be used to
CC generate fusion proteins by linking to the gene for the human
CC immunoglobulin G-Fc portion (SEQID1) for increasing the stability of
CC the fusion protein as compared to the human protein only. The genes and
CC conditions are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated
CC from a range of human tissues disclosed in the specification. The
CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.

| | |
|----|--|
| FT | /partial |
| FT | /product= "Partial Orf 1 protein (AAV58580)" |
| FT | /note= "No initiation codon given in the specification" |
| FT | complement (1900..3171) |
| FT | /*tag- b |
| FT | /product= "Orf 2 protein (AAV58581)" |
| FT | 3415..5556 |
| FT | /*tag- c |
| FT | /product= "Orf 3 protein (AAV58582)" |
| FT | complement (3612..5992) |
| FT | /*tag- d |
| FT | /product= "Orf 4 protein (AAV58583)" |
| FT | 6226..6675 |
| FT | /*tag- e |
| FT | /product= "Orf 5 protein (AAV58584)" |
| FT | 7610..11875 |
| FT | /*tag- f |
| FT | /product= "Type I polyketide synthase, EPOS A |
| FT | 11872..116104 |
| FT | /*tag- g |
| FT | /product= "Non-ribosomal peptide synthetase, EPOS P |
| FT | (AAV58574)" |
| FT | CDS |
| FT | 16251..21749 |
| FT | /*tag- h |
| FT | /product= "Type I polyketide synthase, EPOS B |
| FT | 21746..43519 |
| FT | /*tag- i |
| FT | /product= "Type I polyketide synthase, EPOS C |
| FT | 43524..54920 |
| FT | /*tag- j |
| FT | /product= "Type I polyketide synthase, EPOS D |
| FT | 54935..62254 |
| FT | /*tag- k |
| FT | /product= "Type I polyketide synthase, EPOS E |
| FT | 62369..63628 |
| FT | /*tag- l |
| FT | /product= "Cytochrome P450 oxygenase homologue, EPOS F |
| FT | (AAV58579)" |
| FT | 63779..64333 |
| FT | /*tag- m |
| FT | /product= "Orf 6 protein (AAV58585)" |
| FT | complement (63853..64290) |
| FT | /*tag- n |
| FT | /product= "Orf 7 protein (AAV58586)" |
| FT | 64363..64920 |
| FT | /*tag- o |
| FT | /product= "Orf 8 protein (AAV58587)" |
| FT | complement (64287..64727) |
| FT | /*tag- p |
| FT | /product= "Orf 9 protein (AAV58588)" |
| FT | 65063..65767 |
| FT | /*tag- q |
| FT | /product= "Orf 10 protein" |
| FT | complement (65008..65874) |
| FT | /*tag- r |
| FT | /product= "Orf 11 protein (AAV58590)" |
| FT | complement (65871..66358) |
| FT | /*tag- s |
| FT | /product= "Orf 12 protein (AAV58591)" |
| FT | 66667..67137 |
| FT | /*tag- t |
| FT | /product= "Orf 13 protein (AAV58592)" |
| FT | 67334..68251 |
| FT | /*tag- u |
| FT | /product= "Orf 14 protein (AAV58593)" |
| FT | 68346..68750 |
| FT | /partial |
| FT | /*tag- v |
| FT | /product= "Partial Orf 15 protein (AAV58594)" |
| FT | /note= "No termination codon given in the specification" |
| NN | |
| XX | MO9966028-A2. |

| | | | |
|--------|--|---|---------------------------------|
| PD | | 23-DEC-1999. | |
| XX | PF | 16-JUN-1999; | 99WO-EP04171. |
| XX | PR | 18-JUN-1998; | 98US-0099504. |
| XX | PR | 24-SEP-1998; | 98US-0101631. |
| XX | PR | 05-FEB-1999; | 99US-0118906. |
| PA | (NOVS) | NOVARTIS AG. | |
| PA | (NOVS) | NOVARTIS-ERFINDUNGEN VERW GES MBH. | |
| PI | Schupp T, | Ligon JM, Molnar I, Zirkle R, Goerlach J, Cyr D; | |
| XX | WP: | 2000-097741/08. | |
| DR | P-PSDB: | AAV58573, AAV58574, AAV58575, AAV58576, AAV58577, AAV58578, | |
| DR | | AAV58579, AAV58580, AAV58581, AAV58582, AAV58583, AAV58584, | |
| DR | | AAV58585, AAV58586, AAV58587, AAV58588, AAV58590, AAV58591, | |
| DR | | AAV58592, AAV58593, AAV58594. | |
| PT | New isolated epoethilone synthase genes, used for the recombinant | | |
| PT | production of epoethilone for use in cancer therapy | - | |
| XX | Claim 14; | Page 87-104; 174pp; English. | |
| PS | This sequence represents a 68..75 kb contig from Sorangium cellulosum | | |
| CC | comprising 22 open reading frames (ORFs) and includes genes encoding | | |
| CC | proteins involved in the biosynthesis of epoethilones. Epoethilones A and | | |
| CC | B are 16-membered macrocyclic polyketides with an acylcysteine-derived | | |
| CC | starter unit; polyketides being synthesised from two-carbon building | | |
| CC | blocks," the beta-carbon of which always carries a keto group. Each round | | |
| CC | of two-carbon addition is carried out by a complex of enzymes known as | | |
| CC | the polyketide synthase in a manner similar to fatty acid biosynthesis. | | |
| CC | EPOS A (AAV58573) and EPOS P (AAV58574) are involved in formation of | | |
| CC | the thiazole ring formation of epoethilones, and EPOS B, EPOS C, EPOS D | | |
| CC | and EPOS E (AAV58575-158578) are involved in polyketide backbone | | |
| CC | formation. EPO F (AAV58579) is an epoethilone macrolactone oxidase, and | | |
| CC | the proteins Orl 3 (AAV58582) and Orl14 (AAV58593) are thought to be | | |
| CC | involved in transport. Epoethilones mimic the biological activity of | | |
| CC | taxol, and may be substituted for taxol in cancer chemotherapeutic | | |
| CC | compositions. Epoethilones exhibit a much lower drop in potency against a | | |
| CC | multiply drug-resistant cell line compared with taxol, and are | | |
| CC | considerably less efficiently exported from such cells by the multidrug | | |
| CC | resistance protein (MDR, or P-glycoprotein). Despite the potential of | | |
| CC | epoethilones as anticancer agents, they are problematical to produce on a | | |
| CC | large scale. Epoethilones are too complex for industrial scale chemical | | |
| CC | synthesis, and Sorangium cellulosum is difficult to ferment, producing | | |
| CC | poor yields of epoethilones. The nucleic acids of the invention may be | | |
| CC | used for the recombinant production of epoethilones in a heterologous host | | |
| CC | that is more amenable to fermentation. | | |
| XX | | | |
| SQ | Sequence 68750 BP; 9596 A; 22458 C; 25537 G; 11159 T; 0 other; | | |
| OY | Query Match | 3.0%; | Score 18; DB 21; Length 68750; |
| | Best Local Similarity | 100.0%; | Pred. NO. 11; |
| | Matches 18; Conservative | 0; | Mismatches 0; Indels 0; Gaps 0. |
| | 517 tctgtccacagcgagcgcg 534 | | |
| | | | |
| DG | 19020 ttcgtctcccacgacggcg 19003 | | |
| RESULT | 16 | | |
| ID | AAA29349/C | | |
| XX | AAA29349 standard; DNA; 71989 BP. | | |
| XX | AAA29349; | | |
| DT | 12-SEP-2000 (first entry) | | |
| XX | Sorangium cellulosum epoethilone polyketide synthase operon genomic DNA. | | |
| XX | Epoethilone: polyketide synthase: epok: epob: epoc: epod: epoe: epof: | | |

| Accession | Gene | Protein | Function | Location/Qualifiers |
|-----------|-----------|---|--|---------------------|
| KW | epoK | PA50 epoxidase | ORF_A; ORF_P; promoter; enhancer; anti-fungal; | |
| KX | tubulin | polymerization assay; anti-tumour; cytostatic; ds. | | |
| XX | Sorangium | cellulosum. | | |
| OS | Key | Location/Qualifiers | | |
| PH | Key | 3..992 | | |
| FT | CDS | /*tag= a | | |
| FT | | /label= ORF_A | | |
| FT | | /product= transposase | | |
| FT | | /note= "not part of the PKS" | | |
| FT | | 989..1501 | | |
| FT | CDS | /*tag= b | | |
| FT | | /label= ORF_B | | |
| FT | | /product= transposase | | |
| FT | | /note= "not part of the PKS" | | |
| FT | | 1998..6263 | | |
| FT | CDS | /*tag= c | | |
| FT | | /label= epO_A_gene | | |
| FT | | /note= "encodes the loading domain" | | |
| FT | | 2031..3548 | | |
| FT | misc_RNA | /*tag= d | | |
| FT | | /note= "encodes ketide synthase (KS-Y) of the loading domain" | | |
| FT | | 3621..4661 | | |
| FT | misc_RNA | /*tag= e | | |
| FT | | /note= "encodes acyl transferase (AT) of the loading domain" | | |
| FT | | 4917..5810 | | |
| FT | misc_RNA | /*tag= f | | |
| FT | | /note= "encodes enoyl reductase (ER) of the loading domain, potentially involved in formation of the thiazole moiety" | | |
| FT | | 5856..6155 | | |
| FT | misc_RNA | /*tag= g | | |
| FT | | /note= "encodes acyl carrier protein (ACP) of the loading domain" | | |
| FT | | 6260..10493 | | |
| FT | CDS | /*tag= h | | |
| FT | | /label= epO_B_gene | | |
| FT | | /note= "encodes module 1, the NRPS module" | | |
| FT | | 2031..3548 | | |
| FT | misc_RNA | /*tag= i | | |
| FT | | /note= "encodes condensation domain C2 of the NRPS module" | | |
| FT | | 2031..3548 | | |
| FT | misc_RNA | /*tag= j | | |
| FT | | /note= "encodes condensation domain C2 of the NRPS module" | | |
| FT | | 6861..6887 | | |
| FT | misc_RNA | /*tag= k | | |
| FT | | /note= "encodes heterocyclization signature sequence" | | |
| FT | | 6861..6887 | | |
| FT | misc_RNA | /*tag= l | | |
| FT | | /note= "encodes condensation domain C4 of the NRPS module" | | |
| FT | | 7358..7366 | | |
| FT | misc_RNA | /*tag= m | | |
| FT | | /note= "encodes condensation domain C7 (partial) of the NRPS module" | | |
| FT | | 7898..7921 | | |
| FT | misc_RNA | /*tag= n | | |
| FT | | /note= "encodes adenylation domain A1 of the NRPS module" | | |
| FT | | 7898..7921 | | |
| FT | misc_RNA | /*tag= o | | |
| FT | | /note= "encodes adenylation domain A1 of the NRPS module" | | |
| FT | | 8261..8308 | | |
| FT | misc_RNA | /*tag= p | | |
| FT | | /note= "encodes adenylation domain A3 of the NRPS module" | | |
| FT | | 8411..8422 | | |
| FT | misc_RNA | /*tag= q | | |
| FT | | /note= "encodes adenylation domain A4 of the NRPS module" | | |
| FT | | 8861..8905 | | |

[illegible]

| | | |
|----|----------|----------------------------------|
| FT | | /note= "encodes AT5" |
| FT | misc_RNA | 27966..28574 |
| FT | | /*tag= ao |
| FT | | /note= "encodes DH5" |
| FT | misc_RNA | 29433..30287 |
| FT | | /*tag= ap |
| FT | | /note= "encodes ER5" |
| FT | misc_RNA | 30321..30869 |
| FT | | /*tag= aq |
| FT | | /note= "encodes KR5" |
| FT | misc_RNA | 31077..31373 |
| FT | | /*tag= ar |
| FT | | /note= "encodes ACP5" |
| FT | misc_RNA | 31440..32807 |
| FT | | /*tag= as |
| FT | | /note= "encodes KS6" |
| FT | misc_RNA | 33018..34067 |
| FT | | /*tag= at |
| FT | | /note= "encodes AT6" |
| FT | misc_RNA | 34107..34676 |
| FT | | /*tag= au |
| FT | | /note= "encodes DH6" |
| FT | misc_RNA | 35760..36641 |
| FT | | /*tag= av |
| FT | | /note= "encodes ER6" |
| FT | misc_RNA | 36705..37256 |
| FT | | /*tag= aw |
| FT | | /note= "encodes KR6" |
| FT | misc_RNA | 37470..37769 |
| FT | | /*tag= ax |
| FT | | /note= "encodes ACP6" |
| FT | CDS | 37912..49308 |
| FT | | /*tag= ay |
| FT | | /label= epoe_gene |
| FT | | /note= "encodes modules 7 and 8" |
| FT | misc_RNA | 38014..39375 |
| FT | | /*tag= az |
| FT | | /note= "encodes KS7" |
| FT | misc_RNA | 39589..40626 |
| FT | | /*tag= ba |
| FT | | /note= "encodes AT7" |
| FT | misc_RNA | 41341..41922 |
| FT | | /*tag= bb |
| FT | | /note= "encodes KR7" |
| FT | misc_RNA | 42181..42423 |

Query Match 3.0%; Score 18; DB 21; Length 71989;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY      517 tcgtctccacagcgcgcg 534  
          |||||  
Db      13408 ttcgtctcccacagcgccg 13391
```

RESULT_17
AAx84320
ID AAx84320 standard; DNA; 198 BP.
XX
XX AAx84320;
AC
DT 08-SEP-1999 (first entry)
XX
DE Stealth virus nucleic acid clone, SEQ ID NO: 12.
XX
XX Stealth virus; detection; diagnosis; infection; ss.
OS Stealth virus.
XX
XX W09934019-A1.
PN
FX 08-JUL-1999.
PD
XX

30-DEC-1998; 98WO-US27744.
30-DEC-1997; 97US-0001184.
(MART/) MARTIN W J.
Martin WJ;
WPI: 1999-405521/34.
Novel strains of stealth virus
Claim 19; Page 44; 95pp; English.

This sequence represents a Stealth virus nucleic acid clone. The invention relates to a method of detecting and characterizing a stealth virus by reacting a sample suspected of containing a stealth virus with a probe from a known stealth virus and sequencing the resultant isolated nucleotide. The method comprises the steps of: (a) isolating DNA or RNA from a sample suspected of containing a stealth virus, e.g. a culture of cells showing a viral cytopathic effect; (b) testing the reactivity of the isolated DNA or RNA with a molecular probe that contains at least 18 or more contiguous nucleotides identical to sequence previously identified from a stealth virus; and, optionally (c) sequencing the isolated DNA or RNA molecules that react with the probe. The method is used to detect stealth virus in a biological product, food or in the environment. The method is also used to evaluate agents for their inhibitory or stimulatory effects on stealth virus replication and to determine capacity of the virus to recombine with and potentially alter the nucleic acid sequences of a cell or bacterium.

```

Query Match Similarity      2.9%; Score 17; DB 20; Length 198;
Best Local Similarity      100.0%; Pred. No. 41;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      395  cgacgaggtctacgacg  411
          |||||||
Db       177  cgacgaggtctacgacg  193

RESULT 18
AAC09227/c
ID      AAC09227 standard; cDNA; 353 BP.
XX
AC      AAC09227;
XX
DT      06-OCT-2000 (first entry)
XX
DE      Human secreted protein 5' EST, SEQ ID NO: 13302.
XX
KW      Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX      gene therapy; chromosome mapping; ss.
XX
OS      Homo sapiens.
XX
PN      EP1033401-A2.
XX
PD      06-SEP-2000.
XX
PF      21-FEB-2000; 2000EP-0200610.
XX
PR      26-FEB-1999; 99US-0122487.
XX
PA      (GEST ) GENSET.
XX
PI      Dunas Milne Edwards J, Duclert A, Giordano J;
XX
DR      WPI: 2000-500381/45.
XX
PT      New nucleic acid that is a 5' expressed sequence tag (5' EST) for

```

PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1: SEQ ID 13302; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 353 BP; 94 A; 74 C; 57 G; 127 T; 1 other;

Query Match 2.9%; Score 17; DB 21; Length 353;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 352 aatgcacccaagagag 368
|||||
DB 63 AATGTCACCCCAAGAG 52

RESULT 19
AAC01025
ID AAC01025 standard; cDNA; 386 BP.
XX
AC AAC01025;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 1023.
XX
KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KM gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
PI Dumas Milne Edwards J, Duclert A, Giordano J;
PI MPI: 2000-500381/45.
DR P-PSDB; AAC01019.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1: SEQ ID 1023; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are

CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 386 BP; 125 A; 88 C; 102 G; 70 T; 1 other;

Query Match 2.9%; Score 17; DB 21; Length 386;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 113 aacagcagccagagccaa 129
|||||
DB 92 aacagcagccagagccaa 108

RESULT 20
AAC38084
ID AAC38084 standard; DNA; 513 BP.
XX
AC AAC38084;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 19733.
XX
KM Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143524.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.9%; Score 17; DB 21; Length 513;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 222 tgcgcacccctgcacat 238
Db 481 tgcgcacccctgcacat 497

RESULT 21

AAV16884
ID AAV16884 standard; DNA; 673 BP.

AC AAV16884;

DE 07-AUG-1998 (first entry)

XX Human prostate cancer marker UC Band #31.

XX Prostate cancer; human; marker; diagnosis; treatment; probe; ss.

OS Homo sapiens.

XX MO9804689-A1.

PD 05-FEB-1998.

XX 31-JUL-1996; 96MO-US12516.

XX 31-JUL-1996; 96MO-US12516.

XX (UROC-) UROCOR INC.

XX An G, O'Hara SM, Ralph D, Veltre R;

XX WPI; 1998-130681/12.

XX Human prostate cancer marker - useful for detection and treatment of
XX human prostate cancer

PS Claim 1; Page 135; 229pp; English.

XX This represents a marker sequence for human prostate cancer. Isolated
CC nucleic acid segments shown in AAV16881 to AAV16885, AAV16890 to
CC AAV16903, AAV26351 and AAV26352 which can act as human prostate cancer
CC markers are provided in the specification. It also provides methods for
CC identifying markers for human prostate cancer and for detection of
CC prostate cancer cells. The markers can be identified by amplifying human
CC prostate RNA to provide nucleic acid amplification products, separating
CC the products and identifying those RNA that are differentially expressed
CC between human prostate cancers versus normal or benign human prostate.
CC Prostate cancer cells in a sample can be detected by detecting a nucleic
CC acid in a sample, the nucleic acid being a prostate cancer marker.
CC Primers and probes derived from this marker can be used for the detection
CC of prostate cancer cells in a sample. Antibodies against the protein
CC encoded by the marker nucleic acid fragments, inhibitors of the protein
CC and oligonucleotides antisense to the markers can be used in the
CC treatment of prostate cancer. The antibodies can also be used for the
CC diagnosis of human prostate cancer.

XX Sequence 673 BP; 222 A; 140 C; 145 G; 166 T; 0 other;

Query Match 2.9%; Score 17; DB 19; Length 673;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 537 gaaaccacagagaag 553
Db 542 gaaaccacagagaag 558

RESULT 22

AAZ6019
ID AAZ6019 standard; DNA; 673 BP.

AC AAZ6019;

DE 20-MAY-1999 (first entry)

XX Prostate disease marker gene fragment UC Band #31.

XX Prostate cancer; benign prostatic hyperplasia; marker gene; tumour;

XX differentiation; Reverse Transcription Polymerase Chain Reaction;

XX diagnostic; progression; cancer; metastasis; human; RT-PCR; ss.

OS Homo sapiens.

XX US5882864-A.

PD 16-MAR-1999.

XX 31-JUL-1996; 96US-0692787.

XX 31-JUL-1995; 95US-0001655.

XX 31-JUL-1996; 96US-0692787.

XX (UROC-) UROCOR INC.

XX An G, O'Hara SM, Ralph D, Veltre R;

XX WPI; 1999-214055/18.

XX Diagnosing prostate cancer and benign prostatic hyperplasia cells -
XX using oligonucleotide probes specific for marker genes associated
XX with tumor differentiation and progression in Reverse Transcription
XX Polymerase Chain Reaction analysis

XX Claim 1; Columns 71-72; 74pp; English.

XX The invention relates to methods for diagnosing prostate cancer or benign
CC prostatic hyperplasia cells in a biological sample. The method uses
CC oligonucleotide probes specific for marker genes associated with tumour
CC differentiation and progression in Reverse Transcription Polymerase Chain
CC Reaction (RT-PCR) analysis. The methods are diagnostic techniques useful
CC for detecting and monitoring the progression of benign prostatic
CC hyperplasia and human prostate cancer (the most prevalent form of cancer
CC and a major cause of death in males) prior to the tumor undergoing
CC metastasis, therefore allowing the optimal method of treatment to be
CC determined before the condition becomes life threatening. The present
XX sequence represents a claimed marker gene fragment.

XX Sequence 673 BP; 222 A; 140 C; 145 G; 166 T; 0 other;

Query Match 2.9%; Score 17; DB 20; Length 673;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 537 gaaaccacagagaag 553
Db 542 gaaaccacagagaag 558

RESULT 23

AAZ87504
ID AAZ87504 standard; cDNA; 673 BP.

XX AAZ87504;

DT 19-APR-2000 (first entry)
XX
DE Prostate, breast and bladder cancers detecting biomarker UC Band #31.
XX
XX Nucleic acid marker; biomarker; tumour; prostate cancer; bladder cancer;
KW benign prostatic hyperplasia; BPH; breast cancer; human; immunodetection;
KM diagnosis; ss.
XX
OS Homo sapiens.
XX
PM W09964631-A1.
XX
PD 16-DEC-1999.
XX
PF 11-JUN-1999; 99WO-US13151.
XX
PR 12-JUN-1998; 98US-0097199.
XX
PA (UNOC-) UNOCOR INC.
XX
PI An G, O'Hara SM, Ralph D, Veltri RM;
XX
DR WPI: 2000-116557/10.
XX
PT Novel RNA biomarkers for diagnosis, prognosis and management of
PT prostate, breast and bladder cancer -
XX
XX
PS Claim 1; Page 165; 191pp; English.
XX
XX The invention provides nucleic acid markers of prostate, breast and
CC bladder cancer. The markers are indicators of malignant transformation of
CC prostate, breast and bladder tissues and are diagnostic of the potential
CC for metastatic spread of malignant prostate tumours. The nucleic acid can
CC also be used as targets for therapeutic intervention in prostate cancer,
CC benign prostatic hyperplasia (BPH), bladder cancer or breast cancer. The
CC markers may be used to design specific probes and primers, for the rapid
CC analysis of prostate, bladder or breast biopsy samples. The probes and
CC primers may also be used for in situ hybridization or in situ PCR
CC detection and diagnosis. They may also be used to identify and isolate
CC full length gene sequences from various DNA libraries. Antibodies
CC against the polypeptide products of the markers can be used to treat
CC prostate cancer, bladder cancer or breast cancer. The encoded proteins
CC may be used to detect antibodies. The proteins and antibodies can be
CC used in immunodetection methods for detecting or quantifying the cancers,
CC and for clinical diagnosis of these cancers. The antibodies may also be
CC used for radioimaging to quantify and localize the encoded proteins.
XX
SQ Sequence 673 BP; 222 A; 140 C; 145 G; 166 T; 0 other;

Query Match 2.9%; Score 17; DB 21; Length 673;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 537 gaaccacagcaggaag 553
|||||
DB 542 gaaccacagcaggaag 558

RESULT 24
AAZ33534
ID AAZ33534 standard; cDNA: 697 BP.
XX
AC AAZ33534;
XX
DT 08-DEC-1999 (first entry)
XX
DE Human prostate cancer-associated EST 58.
XX
XX Expressed sequence tag; EST; prostate; tumor; treatment; gene therapy;
KW cancer; tissue specificity; human; ss.
XX
OS Homo sapiens.

XX
PN DE19811194-A1.
XX
PD 16-SEP-1999.
XX
PF 10-MAR-1998; 98DE-1011194.
XX
PR 10-MAR-1998; 98DE-1011194.
XX
PA (META-) METAGEN GES GENOMFORSCHUNG.MBH.
XX
PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
XX
DR WPI: 1999-519629/44.
XX
PD P-PSDB; AAV48440.
XX
PF New nucleic acid expressed at high level in normal prostatic tissue and
PT encoded polypeptides, used to treat cancer and screen for therapeutic
PT agents -
XX
XX
PS Claim 1a; 119; 194pp; German.
XX
XX This invention describes novel nucleic acid sequences (A) that are
CC expressed at high level in normal prostatic tissue. Polypeptides (1)
CC encoded by (A) are used: (a) for identifying agents for treatment of
CC prostatic cancer and (b) for therapy of prostate cancer, optionally
CC where expressed by gene therapy methods. (A) is also used to isolate
CC full-length genes (for gene therapy) and for recombinant production of
CC (1), which can be used to raise specific antibodies. (A) are identified
CC by assembly of ESTs (expressed sequence tags) before these are analyzed
CC for expression pattern (tissue specificity). This approach eliminates
CC many of the false results, as regards tissue specificity, associated
CC with known methods that use single (usually short) ESTs. AAZ33477-233540
CC represent expressed sequence tags described in the method of the
CC invention.
XX
SQ Sequence 697 BP; 223 A; 140 C; 196 G; 138 T; 0 other;

Query Match 2.9%; Score 17; DB 20; Length 697;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 113 aacagcagccagccaa 129
|||||
DB 110 aacagcagccagccaa 126

RESULT 25
AAC39218
ID AAC39218 standard; DNA: 763 BP.
XX
AC AAC39218;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 23810.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PM EP1033405-A2.
XX
PN
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0129845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130049.
PR 23-APR-1999; 9905-0130510.
PR 28-APR-1999; 9905-0130891.
PR 30-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 30-APR-1999; 9905-0132407.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 07-MAY-1999; 9905-0132487.
PR 11-MAY-1999; 9905-0132863.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134321.
PR 18-MAY-1999; 9905-0134370.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138340.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139452.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.

PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0158295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.9%; Score 17; DB 21; Length 763;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 217 gccgctgctgcatccccc 233
DB 204 gccgctgctgcatccccc 220

RESULT 26
AAC99839/C
ID AAC99839 standard; cDNA; 800 BP.

AC AAC99839;
DT 13-MAR-2001 (first entry)
XX

DE Human secreted protein gene 22 SEQ ID NO:32.

XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KM antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic;
KM cerebroprotective; neuroprotective; antibacterial; virucide;
KM fungicide; ophthalmological; gene therapy; pathological condition;
KM autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;
KM neoplasm; cardiovascular disorder; cardiac arrest; cerebral ischaemia;
KM cerebrovascular disorder; angiogenesis; nervous system disorder;
KM Alzheimer's disease; infection; ocular disorder; corneal infection;
KM wound healing; skin aging; food additive; preservative;
KM chromosome 11; ss.

OS Homo sapiens.

PN WO200070042-A1.

PD 23-NOV-2000.

PF 11-MAY-2000; 2000MO-US12788.

PR 13-MAY-1999; 99US-0134068.

PA (HDMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Moore PA, Young PE, Komatsoulis GA, Birse CE;
PI Duan RD, Florence KA, Soppet DR;
XX WPI. 2000-679828/66.
DR P-PSDB; AAB56098.
XX
PT Isolated nucleic acid molecule encoding a human secreted protein is
PS used in preventing, treating or ameliorating a medical condition -
XX Claim 1; Page 868; 1065pp; English.

The polynucleotide sequences given in AAC99818 to AAC99977 encode the human secreted proteins given in AAB56077 to AAB56362. Human secreted proteins have activities based on the tissues and cells the genes are expressed in. Examples of activities include: immunosuppressive; antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective; neuroprotective; antibacterial; virucide; fungicide; and ophthalmological. The human secreted polynucleotides and proteins can be used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The proteins can also be used as a food additive or preservative to increase or decrease storage capabilities. AAC99809 to AAC99817 and AAB5076 represent sequences used in the exemplification of the present invention.

Sequence 800 BP; 154 A; 187 C; 154 G; 305 T; 0 other;

Query Match 2.9%; Score 17; DB 21; Length 800;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 aacagcagccagcccaa 129
DB 743 AACAGCAGCCAGGCCAA 727

RESULT 27

ID AAC42507 standard; DNA; 1148 BP.

AC AAC42507;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 35818.

XX Hybridisation assay; genetic mapping; gene expression control;

XX protein identification; signal transduction pathway;

XX metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128224.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-01300449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132407.
PR 30-APR-1999; 99US-0132408.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140655.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147203.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.

PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0156029.
PR 08-OCT-1999; 99US-0156232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.9%; Score 17; DB 21; Length 1148;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 222 tgcgcacccctgcacat 238
|||||
Db 593 tgcgcacccctgcacat 609

RESULT 28

AAC49984
ID AAC49984 standard; DNA; 1273 BP.

AC AAC49984;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 63167.

KW Hybridisation assay; genetic mapping; gene expression control;

KW Protein identification; signal transduction pathway;

KW Metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 23-MAR-1999; 99US-0123548.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 15-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156566.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.

PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.9%; Score 17; DB 21; Length 1273;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 222 tgcgcacccctgcacat 238
|||||
Db 522 tgcgcacccctgcacat 538

RESULT 29
AAC37359
ID AAC37359 standard; DNA: 1277 BP.
XX AC AAC37359;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 17090.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129645.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.

| | | |
|-------|--------------|---------------|
| PR | 30-APR-1999; | 99US-0132407. |
| PR | 04-MAY-1999; | 99US-0132484. |
| PR | 05-MAY-1999; | 99US-0132485. |
| PR | 06-MAY-1999; | 99US-0132486. |
| PR | 06-MAY-1999; | 99US-0132487. |
| PR | 07-MAY-1999; | 99US-0132863. |
| PR | 11-MAY-1999; | 99US-0134256. |
| PR | 14-MAY-1999; | 99US-0134218. |
| PR | 14-MAY-1999; | 99US-0134219. |
| PR | 14-MAY-1999; | 99US-0134221. |
| PR | 14-MAY-1999; | 99US-0134370. |
| PR | 18-MAY-1999; | 99US-0134768. |
| PR | 19-MAY-1999; | 99US-0134941. |
| PR | 20-MAY-1999; | 99US-0135124. |
| PR | 21-MAY-1999; | 99US-0135353. |
| PR | 24-MAY-1999; | 99US-0135629. |
| PR | 25-MAY-1999; | 99US-0136021. |
| PR | 27-MAY-1999; | 99US-0136392. |
| PR | 28-MAY-1999; | 99US-0136782. |
| PR | 01-JUN-1999; | 99US-0137222. |
| PR | 03-JUN-1999; | 99US-0137528. |
| PR | 04-JUN-1999; | 99US-0137502. |
| PR | 07-JUN-1999; | 99US-0137724. |
| PR | 08-JUN-1999; | 99US-0138094. |
| PR | 10-JUN-1999; | 99US-0138540. |
| PR | 14-JUN-1999; | 99US-0138847. |
| PR | 16-JUN-1999; | 99US-0139119. |
| PR | 16-JUN-1999; | 99US-0139452. |
| PR | 17-JUN-1999; | 99US-0139453. |
| PR | 18-JUN-1999; | 99US-0139492. |
| PR | 18-JUN-1999; | 99US-0139454. |
| PR | 18-JUN-1999; | 99US-0139455. |
| PR | 18-JUN-1999; | 99US-0139456. |
| PR | 18-JUN-1999; | 99US-0139457. |
| PR | 18-JUN-1999; | 99US-0139458. |
| PR | 18-JUN-1999; | 99US-0139459. |
| PR | 18-JUN-1999; | 99US-0139460. |
| PR | 18-JUN-1999; | 99US-0139461. |
| PR | 18-JUN-1999; | 99US-0139462. |
| PR | 18-JUN-1999; | 99US-0139463. |
| PR | 18-JUN-1999; | 99US-0139750. |
| PR | 21-JUN-1999; | 99US-0139763. |
| PR | 22-JUN-1999; | 99US-0139817. |
| PR | 23-JUN-1999; | 99US-0139899. |
| PR | 23-JUN-1999; | 99US-0140353. |
| PR | 24-JUN-1999; | 99US-0140354. |
| PR | 28-JUN-1999; | 99US-0140695. |
| PR | 29-JUN-1999; | 99US-0140823. |
| PR | 30-JUN-1999; | 99US-0140991. |
| PR | 01-JUL-1999; | 99US-0141287. |
| PR | 01-JUL-1999; | 99US-0141842. |
| PR | 02-JUL-1999; | 99US-0142134. |
| PR | 06-JUL-1999; | 99US-0142055. |
| PR | 08-JUL-1999; | 99US-0142390. |
| PR | 09-JUL-1999; | 99US-0142803. |
| PR | 12-JUL-1999; | 99US-0142920. |
| PR | 13-JUL-1999; | 99US-0142977. |
| PR | 14-JUL-1999; | 99US-0143542. |
| PR | 15-JUL-1999; | 99US-0143624. |
| PR | 16-JUL-1999; | 99US-0144005. |
| PR | 16-JUL-1999; | 99US-0144085. |
| PR | 16-JUL-1999; | 99US-0144086. |
| PR | 19-JUL-1999; | 99US-0144325. |
| PR | 19-JUL-1999; | 99US-0144331. |
| PR | 19-JUL-1999; | 99US-0144332. |
| PR | 19-JUL-1999; | 99US-0144333. |
| PR | 19-JUL-1999; | 99US-0144334. |
| PR | 19-JUL-1999; | 99US-0144335. |
| PR | 20-JUL-1999; | 99US-0144352. |
| PR | 20-JUL-1999; | 99US-0144632. |
| PR | 21-JUL-1999; | 99US-0144884. |
| PR | 21-JUL-1999; | 99US-0145086. |
| <hr/> | | |
| PR | 21-JUL-1999; | 99US-0145088. |
| PR | 22-JUL-1999; | 99US-0145085. |
| PR | 22-JUL-1999; | 99US-0145087. |
| PR | 22-JUL-1999; | 99US-0145089. |
| PR | 22-JUL-1999; | 99US-0145192. |
| PR | 23-JUL-1999; | 99US-0145145. |
| PR | 23-JUL-1999; | 99US-0145218. |
| PR | 26-JUL-1999; | 99US-0145224. |
| PR | 27-JUL-1999; | 99US-0145276. |
| PR | 27-JUL-1999; | 99US-0145913. |
| PR | 27-JUL-1999; | 99US-0145918. |
| PR | 28-JUL-1999; | 99US-0145919. |
| PR | 02-AUG-1999; | 99US-0145951. |
| PR | 02-AUG-1999; | 99US-0146386. |
| PR | 02-AUG-1999; | 99US-0146388. |
| PR | 03-AUG-1999; | 99US-0146389. |
| PR | 04-AUG-1999; | 99US-0147038. |
| PR | 04-AUG-1999; | 99US-0147204. |
| PR | 05-AUG-1999; | 99US-0147302. |
| PR | 05-AUG-1999; | 99US-0147192. |
| PR | 06-AUG-1999; | 99US-0147260. |
| PR | 06-AUG-1999; | 99US-0147303. |
| PR | 09-AUG-1999; | 99US-0147416. |
| PR | 09-AUG-1999; | 99US-0147493. |
| PR | 10-AUG-1999; | 99US-0147935. |
| PR | 11-AUG-1999; | 99US-0148171. |
| PR | 12-AUG-1999; | 99US-0148319. |
| PR | 13-AUG-1999; | 99US-0148341. |
| PR | 13-AUG-1999; | 99US-0148565. |
| PR | 16-AUG-1999; | 99US-0148684. |
| PR | 17-AUG-1999; | 99US-0149368. |
| PR | 18-AUG-1999; | 99US-0149175. |
| PR | 20-AUG-1999; | 99US-0149426. |
| PR | 20-AUG-1999; | 99US-0149722. |
| PR | 20-AUG-1999; | 99US-0149723. |
| PR | 23-AUG-1999; | 99US-0149929. |
| PR | 23-AUG-1999; | 99US-0149902. |
| PR | 23-AUG-1999; | 99US-0149330. |
| PR | 25-AUG-1999; | 99US-0150566. |
| PR | 26-AUG-1999; | 99US-0150884. |
| PR | 27-AUG-1999; | 99US-0151065. |
| PR | 27-AUG-1999; | 99US-0151066. |
| PR | 27-AUG-1999; | 99US-0151080. |
| PR | 30-AUG-1999; | 99US-0151303. |
| PR | 31-AUG-1999; | 99US-0151438. |
| PR | 01-SEP-1999; | 99US-0151930. |
| PR | 07-SEP-1999; | 99US-0152363. |
| PR | 10-SEP-1999; | 99US-0153070. |
| PR | 13-SEP-1999; | 99US-0153758. |
| PR | 15-SEP-1999; | 99US-0154018. |
| PR | 16-SEP-1999; | 99US-0154039. |
| PR | 20-SEP-1999; | 99US-0154779. |
| PR | 22-SEP-1999; | 99US-0155139. |
| PR | 23-SEP-1999; | 99US-0155486. |
| PR | 24-SEP-1999; | 99US-0155659. |
| PR | 28-SEP-1999; | 99US-0155658. |
| PR | 29-SEP-1999; | 99US-0156596. |
| PR | 04-OCT-1999; | 99US-0157117. |
| PR | 05-OCT-1999; | 99US-0157753. |
| PR | 06-OCT-1999; | 99US-0157865. |
| PR | 07-OCT-1999; | 99US-0158029. |
| PR | 08-OCT-1999; | 99US-0158232. |
| PR | 12-OCT-1999; | 99US-0158393. |
| PR | 13-OCT-1999; | 99US-0158393. |
| PR | 13-OCT-1999; | 99US-0159293. |
| PR | 13-OCT-1999; | 99US-0159294. |
| PR | 14-OCT-1999; | 99US-0159329. |
| PR | 14-OCT-1999; | 99US-0159330. |
| PR | 14-OCT-1999; | 99US-0159331. |
| PR | 14-OCT-1999; | 99US-0159637. |
| PR | 14-OCT-1999; | 99US-0159638. |
| PR | 18-OCT-1999; | 99US-0159584. |
| PR | 21-OCT-1999; | 99US-0160741. |

PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.9%; Score 17; DB 21; Length 1277;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 222 tgcgcacccctgcac 238
|||||
DB 524 tgcgcacccctgcac 540

RESULT 30
AAV64526/c
ID AAV64526 standard; DNA: 1338 BP.
XX
AC AAV64526;

DT 27-JAN-1999 (first entry)
XX

DE M. tuberculosis immunogenic polypeptide XP31 DNA.

KM Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
KW vaccine; pharmaceutical; infection; diagnosis; ss.

OS Mycobacterium tuberculosis.

PN W09816646-A2.

PD 23-APR-1998.

PF 07-OCT-1997; 97WO-US18293.

PR 13-MAR-1997; 97US-0818112.

PR 11-OCT-1996; 96US-0730510.

PA (CORI-) CORIXA CORP.

PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ;
Red SG, Skeiky YAW, Twardzik DR, Vedvick TS;

DR WPI; 1998-261042/23.

PT Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used
to develop products for the detection of M. tuberculosis infection
and for diagnosis, treatment and prevention of tuberculosis
Claim 31a: Page 162-163; 230pp; English.

XX This sequence encodes an immunogenic portion of a soluble Mycobacterium
tuberculosis (MT) antigen which can be used in a method for inducing
CC protective immunity against tuberculosis (TB). This sequence can be
CC formulated into vaccines and/or pharmaceutical compositions for
CC immunizing against M. tuberculosis infection or may be used for the
CC diagnosis of tuberculosis.

XX Sequence 1338 BP; 199 A; 432 C; 485 G; 222 T; 0 other;

Query Match 2.9%; Score 17; DB 19; Length 1338;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 505 gccagcgagcctgcac 521
|||||
DB 1147 GCCAGCGGCGCTGTC 1131

RESULT 31
AAV44417/c
ID AAV44417 standard; DNA: 1338 BP.
XX
AC AAV44417;

DT 09-NOV-1998 (first entry)
XX

DE Mycobacterium tuberculosis antigen XP31 DNA.

KM Tuberculosis; infection; diagnosis; antigen; XP31; ss.

OS Mycobacterium tuberculosis strain Erdman.

PN W09816645-A2.

PD 23-APR-1998.

PF 07-OCT-1997; 97WO-US18214.

PR 13-MAR-1997; 97US-0818111.

PR 11-OCT-1996; 96US-0729622.

PA (CORI-) CORIXA CORP.

PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ;
Red SG, Skeiky YAW, Twardzik DR, Vedvick TS;

DR WPI; 1998-251292/22.

PT New isolated Mycobacterium tuberculosis polypeptides and DNA - used
to develop products for the detection of M. tuberculosis infection
and diagnosis of tuberculosis

XX Claim 11a: Page 171-172; 250pp; English.

XX This DNA sequence codes for a portion of Mycobacterium tuberculosis
CC antigen XP31. It was isolated from a M. tuberculosis strain Erdman
CC genomic DNA expression library using sera from patients having
CC extrapulmonary tuberculosis. The clone bears some similarity to
CC known sequences. The invention relates to compositions and methods
CC for diagnosing tuberculosis. It provides polypeptides (see
CC AAM64291-W64379) comprising an antigenic portion of a soluble M.
CC tuberculosis antigen, or an immunogenic portion of an M.
CC tuberculosis antigen, as well as DNA sequences encoding such
CC polypeptides, recombinant expression vectors and transformed or
CC transfected host cells. Also claimed are methods and diagnostic
CC kits for detecting M. tuberculosis infection in a patient using
CC these polypeptides, antibodies or oligonucleotide probes and
CC primers.

XX Sequence 1338 BP; 199 A; 432 C; 485 G; 222 T; 0 other;

Query Match 2.9%; Score 17; DB 19; Length 1338;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 505 gccagcgagcctgcac 521
|||||
DB 1147 GCCAGCGGCGCTGTC 1131

RESULT 32
AA219327/C
ID AA219327 standard; DNA; 1338 BP.
XX
AC AA219327;
XX
DT 05-NOV-1999 (first entry)
XX
DE M. tuberculosis antigen XP31 DNA sequence.
XX
KW Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
KW immunotherapy; diagnosis; immunisation; vaccine; infection;
KW immune response; skin test; ss.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9942076-A2.
XX
PD 26-AUG-1999.
XX
PE 17-FEB-1999; 99WO-US03268.
XX
PR 05-MAY-1998; 98US-0072967.
PR 18-FEB-1998; 98US-0025197.
XX
PA (CORI-) CORIXA CORP.
XX
PI Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;
PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
XX
DR WPI; 1999-527409/44.
XX
PT New antigens from Mycobacterium tuberculosis useful in diagnostic
PT skin tests and protective or therapeutic vaccines or compositions
XX
PS Claim 11; Page 157-158; 299pp; English.
XX
CC The present invention describes polypeptides comprising an immunogenic
CC part of a Mycobacterium tuberculosis antigen (Ag). Also described
CC are vaccines and fusion protein containing M. tuberculosis Ag's.
CC M. tuberculosis Ag's, DNAs encoding them, derived fusion proteins and
CC other polypeptides fragments, can be used in pharmaceutical compositions
CC or vaccines to generate a protective or therapeutic immune response to
CC M. tuberculosis and as reagents in skin tests for diagnosis of
CC tuberculosis. Ag can induce proliferation of, or cytokine secretion
CC by, T, B or natural killer cells and/or macrophages in
CC tuberculosis-immune subjects. AA219249 to AA219460 and AA219083 to
CC AA219325 are used in the exemplification of the present invention.
XX
SQ Sequence 1338 BP; 199 A; 432 C; 485 G; 222 T; 0 other;

Query Match 2.9%; Score 17; DB 20; Length 1338;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 505 gccagcgcgctgcgtc 521
|||||
Db 1147 GCCAGCGCGCTCGTC 1131

RESULT 33
AA219115/C
ID AA219115 standard; DNA; 1338 BP.
XX
AC AA219115;
XX
DT 05-NOV-1999 (first entry)
XX
DE M. tuberculosis recombinant antigen DNA encoding XP31.
XX
KW Antigen; diagnosis; detection; infection; antibody; immunisation;

KW vaccine; immunity; ss.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9942118-A2.
XX
PD 26-AUG-1999.
XX
PE 17-FEB-1999; 99WO-US03265.
XX
PR 05-MAY-1998; 98US-0072596.
PR 18-FEB-1998; 98US-0024753.
XX
PA (CORI-) CORIXA CORP.
XX
PI Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;
PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
XX
DR WPI; 1999-527416/44.
XX
PT New polypeptide comprising antigenic portions of M. tuberculosis
XX
PS Claim 11a; Page 202-203; 323pp; English.
XX
CC This invention describes novel recombinant antigens and their encoding
CC nucleic acids derived from Mycobacterium tuberculosis. The novel
CC polypeptides are useful for detecting M. tuberculosis infection in a
CC biological sample by detecting antibodies which bind with the
CC polypeptides, and are useful as vaccines for immunizing against
CC M. tuberculosis infection. The new detection methods are needed as
CC current vaccination strategies do not provide 100% immunity.
XX
SQ Sequence 1338 BP; 199 A; 432 C; 485 G; 222 T; 0 other;

Query Match 2.9%; Score 17; DB 20; Length 1338;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 505 gccagcgcgctgcgtc 521
|||||
Db 1147 GCCAGCGCGCTCGTC 1131

RESULT 34
AAC79935
ID AAC79935 standard; cDNA; 1647 BP.
XX
AC AAC79935;
XX
DT 12-FEB-2001 (first entry)
XX
DE Human secreted protein encoding cDNA for gene 38.
XX
KW Secreted protein; human; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW neurotropic; neuroprotective; antibacterial; virucide; fungicide;
KW ophthalmological; gene therapy; treatment; autoimmune disease; infection;
KW cardiovascular disorder; nervous system disorder; ocular disorder;
KW wound healing; epithelial cell proliferation; skin aging; mental state;
KW transplantation; metabolism modulation; ss.
XX
OS Homo sapiens.
XX
PN WO200055200-A1.
XX
PD 21-SEP-2000.
XX
PE 09-MAR-2000; 2000WO-US06042.
XX
PR 12-MAR-1999; 99US-0124143.
PR 03-DEC-1999; 99US-0168663.
XX

| | | |
|--------------------------|---|--------------------------|
| PA | (HUMA-) | HUMAN GENOME SCI INC. |
| PX | | |
| PI | Rosen CA, | Ruben SM, Komatsoulis G; |
| DR | WPI: | 2000-656008/63. |
| DR | P-P5DB; | AAB44954. |
| PX | | |
| PT | Isolated human secretory proteins, nucleic acids encoding them and | |
| PR | antibodies directed against them, useful for diagnosing and treating | |
| PT | disorders related to the proteins such as cancer, Alzheimer's disease | |
| PT | and Parkinsons - | |
| PS | Claim 1a; Page 370-371; 453pp: | English. |
| PX | | |
| CC | This invention describes a novel isolated polypeptide (I) and its | |
| CC | encoding nucleic acid molecule (II) which have immunosuppressive, | |
| CC | antiarthritic, antirheumatic, antiproliferative, cytostatic, cardiant, | |
| CC | vasotropic, cerebroprotective, neurotropic, neuroprotective, antibacterial, | |
| CC | viralicide, fungicide and ophthalmological activity and which can be used | |
| CC | for gene therapy. (I) and (II) are used to prevent, treat or ameliorate | |
| CC | a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, | |
| CC | dogs, chickens or sheep. (I) and (II) are also used in diagnosing a | |
| CC | pathological condition or susceptibility to a pathological condition. The | |
| CC | antibodies to (II) can also be used in alleviating symptoms associated | |
| CC | with the disorders and in diagnostic immunoassays e.g. radioimmunoassays | |
| CC | or enzyme linked immunosorbent assays (ELISA). Disorders which are | |
| CC | diagnosed or treated include autoimmune diseases e.g. rheumatoid | |
| CC | arthritis, hyperproliferative disorders e.g. neoplasms of the breast or | |
| CC | liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular | |
| CC | disorders e.g. cerebral ischemia, angiodenesis, nervous system disorders | |
| CC | e.g. Alzheimer's disease, infections caused by bacteria, viruses and | |
| CC | fungal and ocular disorders e.g. corneal infection. The polypeptides can | |
| CC | also be used to aid wound healing and epithelial cell proliferation, to | |
| CC | prevent skin aging due to sunburn, to maintain organs before | |
| CC | transplantation, for supporting cell culture of primary tissues, to | |
| CC | regenerate tissues and in chemotaxis. The polypeptides are used to | |
| CC | modulate mammalian metabolism, to change mammal's mental state or | |
| CC | physical state by influencing biohythms circadian rhythms, depression | |
| CC | tendency for violence tolerance for pain, reproductive capabilities, | |
| CC | hormonal or endocrine levels, appetite, libido, memory, stress or other | |
| CC | cognitive qualities, as a food additive or preservative, such as to | |
| CC | increase or decrease storage capabilities, fat content, lipid, protein, | |
| CC | carbohydrates, vitamins, minerals, cofactors or other nutritional | |
| CC | components. | |
| SQ | Sequence 1647 BP; 425 A; 301 C; 329 G; 583 T; 9 other: | |
| XX | | |
| XX | | |
| Query Match | 2.9%; Score 17; DB 21; Length 1647; | |
| Best Local Similarity | 100.0%; Pred. No. 39; | |
| Matches 17; Conservative | 0; Mismatches 0; Indels 0; Gaps 0; | |
| OY | 485 tgggtgcagctaataagt 501 | |
| | | |
| Db | 1018 tgggtgcagctaataagt 1034 | |
| RESULT 35 | | |
| AAAT34620 | standard; DNA; 3337 BP. | |
| XX | | |
| AC | AAAT34620; | |
| XX | | |
| DT | 12-NOV-1996 (first entry) | |
| XX | | |
| DE | P. vivax ESP-1 blood stage antigen coding sequence. | |
| XX | | |
| KM | ESP-1: blood stage antigen; diagnosis; malaria; infection; | |
| KM | causative agent; antibody; monoclonal; polyclonal; assay; ds. | |
| XX | | |
| OS | Plasmodium vivax (clone PWMB3.3.1). | |
| XX | | |
| FH | Key Location/Qualifiers | |

| | | |
|-----------|--|---|
| FT | Exon | 1..91 |
| FT | /tag= a | "encodes initial (N-terminal) sequence of |
| FT | /note= | hydrophobic amino acids" |
| FT | Intron | 92..230 |
| FT | /tag= b | "contains typical malaria intervening |
| FT | /note= | sequence splice sites splice sites" |
| FT | exon | 231..3197 |
| FT | /tag= c | |
| PX | US5532133-A. | |
| XX | | |
| XX | 02-JUL-1996. | |
| XX | | |
| PF | 02-JUN-1993; | 93US-0072610. |
| XX | | |
| PR | 02-JUN-1993; | 93US-0072610. |
| XX | | |
| PA | (UYNY) UNIV NEW YORK STATE. | |
| PI | Barnwell JM; | |
| XX | | |
| DR | WPI; 1996-321110/32. | |
| XX | P-PSDB; AAR98747. | |
| PT | Antibodies to Plasmodium vivax blood stage antigens - used to | |
| PT | diagnose malaria and to determine whether P. vivax is the species | |
| PT | responsible for infection | |
| PS | Example 4; Column 15-20; 22pp; English. | |
| XX | | |
| CC | The present sequence encodes a species-specific Plasmodium vivax | |
| CC | malarial antigen, PvESP-1. The gene appears to be missing a small | |
| CC | portion of its 5' end. This protein is secreted into the plasma of | |
| CC | a susceptible mammalian host after infection. Monoclonal/polyclonal | |
| CC | antibodies can be utilised in assays used to diagnose malaria, as well | |
| CC | as to determine whether P. vivax is the species responsible for the | |
| CC | infection. | |
| XX | | |
| SQ | Sequence 3337 BP; 1304 A; 467 C; 875 G; 691 T; 0 other; | |
| Oy | Query Match | 2.9%; Score 17; DB 17; Length 3337; |
| | Best Local Similarity | 100.0%; Pred. No. 38; |
| | Matches 17; Conservative | 0; Mismatches 0; Indels 0; Gaps 0; |
| Db | 539 aaccacagcagaagaaga 555 | |
| | 2029 aaccacagcagaagaaga 2045 | |
| RESULT 36 | | |
| AAAX15174 | | |
| ID | AAAX15174 standard; DNA; 3337 BP. | |
| XX | | |
| AC | AAAX15174; | |
| XX | | |
| DT | 28-APR-1999 (first entry) | |
| DE | DNA encoding a secreted blood-stage protein called PvESP-1. | |
| XX | | |
| KM | Erythrocyte secreted protein-1; PvESP-1; malarial antigen; | |
| KM | blood-stage protein; malaria; monoclonal antibody 1D11G10; ds. | |
| OS | Plasmodium vivax. | |
| XX | | |
| FH | Key | Location/Qualifiers |
| FT | CDS | 1..3197 |
| FT | /tag= a | |
| FT | /note= "contains 1 intron" | |
| FT | 1..91 | |
| FT | /tag= b | |

```
FT      /number= 1
FT      92..230
FT      /*tag= c
FT      /number= 1
FT      exon      231..3194
FT      /*tag= d
FT      /number= 2
XX
XX      US5874527-A.
XX
XX      23-FEB-1999.
XX
XX      30-SEP-1996; 96US-0719822.
XX
XX      02-JUN-1993; 93US-0072610.
XX      07-JUN-1995; 95US-0478417.
XX      30-SEP-1996; 96US-0719822.
XX
XX      (UYNV ) UNIV NEW YORK STATE.
XX
XX      Barnwell JM;
XX
XX      WPI; 1999-180063/15.
XX      P-PSDB; AAW97039.
XX
XX      Plasmodium vivax peptide antigen - for diagnosis of malaria caused
XX      by Plasmodium vivax
XX
XX      Example 4; Fig 5A-C; 23pp; English.
XX
XX      The present sequence encodes a C-terminal erythrocyte secreted
XX      protein-1 (PvESP-1) of Plasmodium vivax. PvESP-1 is a malarial
XX      antigen which is a secreted blood-stage protein present in detectable
XX      amounts in biological samples from individuals infected with P. vivax.
XX      The protein comprises an epitope not present in other Plasmodium species
XX      that cause malaria in humans, and is bound by monoclonal antibody
XX      1D1G10. The peptide antigen can be used in immunoassays for diagnosis
XX      of malaria caused by P. vivax and/or can be used to produce antibodies
XX      for use in such immunoassays.
XX
XX      Sequence 3337 BP; 1304 A; 467 C; 875 G; 691 T; 0 other;
SQ
Query Match      2.9%; Score 17; DB 20; Length 3337;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      539 aaccaacagagaaga 555
          |||||||||
DB      2029 aaccaacagagaaga 2045

RESULT 37
AAC74895
ID      AAC74895 standard; cDNA; 6171 BP.
XX
XX      AAC74895;
AC
XX      08-FEB-2001 (first entry)
DT
XX
XX      Human ORF450 polynucleotide sequence SEQ ID NO:899.
DE
XX
XX      Human; open reading frame; ORF; detection; cytosolic; hepatotropic;
XX      vulnery; antiparasitic; antiparkinsonian; neurotropic; neuroprotective;
XX      anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
XX      immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX      hypotensive; dermatological; immunosuppressive; antiinflammatory;
XX      antiviral; antibacterial; antifungal; antineumatic; antihyroid;
XX      antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
XX      neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX      cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
XX      cholesterol ester storage; systemic lupus erythematosus; infection;
XX      severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
```

```
KW      allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW      bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW      thrombosis; contraceptive; ss.
XX
XX      Homo sapiens.
XX
XX      WO200058473-A2.
XX
XX      05-OCT-2000.
XX
XX      31-MAR-2000; 2000WO-US08621.
XX
XX      31-MAR-1999; 99US-0127607.
XX      02-APR-1999; 99US-0127636.
XX      05-APR-1999; 99US-0127728.
XX      30-MAR-2000; 2000US-0540763.
XX
XX      (CURA-) CURAGEN CORP.
XX
XX      Shimkets RA, Leach M;
XX
XX      WPI; 2000-602362/57.
XX      P-PSDB; AAB40686.
XX
XX      Novel nucleic acids and peptides derived from open reading frame X,
XX      useful for treating e.g. cancers, proliferative disorders,
XX      neurodegenerative disorders and cardiovascular disease -
XX
XX      Claim 5; Page 910-914; 5507pp; English.
XX
XX      AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX      which represent the human ORF open reading frames 1 to 3161. The ORF
XX      sequences have activities such as: cytosolic; hepatotropic; vulnery;
XX      antiparasitic; antiparkinsonian; neurotropic; neuroprotective;
XX      osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
XX      immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
XX      antidiabetic; hypotensive; dermatological; immunosuppressive;
XX      antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;
XX      antihyroid; and antianaemic. The sequences can be used for determining
XX      the presence of or predisposition to, or preventing or treating
XX      pathological conditions associated with an ORF-associated disorder. The
XX      nucleic acids can be used to express ORF proteins in gene therapy
XX      vectors. The proteins and nucleic acids may be used to treat cancers,
XX      proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX      graft vs host disease, cardiovascular disease, diabetes mellitus,
XX      hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX      erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX      bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX      allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX      nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX      coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX      Sequence 6171 BP; 1659 A; 1317 C; 1436 G; 1759 T; 0 other;
SQ
Query Match      2.9%; Score 17; DB 21; Length 6171;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      245 ctgttcagcgccagcc 261
          |||||||||
DB      1615 ctgttcagcgccagcc 1631

RESULT 38
AAA58762
ID      AAA58762 standard; DNA; 11212 BP.
XX
XX      AAA58762;
AC
XX      20-OCT-2000 (first entry)
DT
XX
XX      DNA encoding a cyclohexadepsipeptide synthetase.
DE
```

XX Cyclohexadepsipeptide synthetase; filamentous fungal cell;
KM cyclohexadepsipeptide; antibiotic; ss.
XX
OS Fusarium venenatum.
XX
FH Key Location/Qualifiers
FT CDS 270..9659
FT /tag- a
FT /product- "cyclohexadepsipeptide synthetase"
XX
XX MO200042203-A2.
XX
XX 20-JUL-2000.
XX
XX 13-JAN-2000; 2000MO-US00913.
XX
XX 13-JAN-1999; 99US-0229862.
XX
XX (NOVO) NOVO NORDISK BIOTECH INC.
XX
XX Berka RM, Rey MM, Yoder WT;
XX
XX MPI; 2000-482633/42.
XX
XX P-PSDB; AAB07427.
XX
XX Producing a heterologous polypeptide for production of antibiotics
PT comprising a mutant of a parent filamentous fungal cell
PT comprising a nucleic acid sequence encoding cyclohexadepsipeptide -
XX
XX Claim 55; Flg 1A-I; 76pp; English.
XX
XX The present sequence encodes a cyclohexadepsipeptide synthetase
CC polypeptide. The specification describes a method for producing a
CC heterologous polypeptide. The method comprises cultivating a
CC mutant of a parent filamentous fungal cell, which produces less
CC cyclohexadepsipeptide than the parent filamentous fungal cell when
CC cultured under the same conditions. The method if used for the
CC production of biologically active compounds e.g. antibiotics.
XX
SQ Sequence 11212 BP; 2808 A; 2844 C; 2658 G; 2901 T; 1 other;

Query Match 2.9%; Score 17; DB 21; Length 11212;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 ctgcgcgtcgtcatcc 231
|||||
DB 697 ctgcgcgtcgtcatcc 713

RESULT 39
AAD01009/c
ID AAD01009 standard; DNA; 11672 BP.
XX
XX AAD01009;
XX
XX 21-SEP-2000 (first entry)
XX
XX Escherichia coli phn operon sequence #2.
XX
XX Phosphonate herbicide tolerance; aminomethyl phosphonic acid; AMPA; AAT;
KM acyltransferase; transacylase; recombinant plant; expression cassette;
KM corn; tobacco; wheat; cotton; canola; rice; chloroplast transit peptide;
KM CTP; glyphosate oxidase; GOX; glyphosate oxidoreductase; phn operon;
KM self-fertilisation; hetero-fertilisation; ds.
XX
XX Escherichia coli.
XX
OS
XX
FH Key Location/Qualifiers
FT CDS 9726..10160
FT /tag- a

FT /product- "Escherichia coli AMPA acyltransferase (AAT)"
FT /function- "Transfers acyl group from an acylcarrier
FT (CoA) to the free amino group of aminomethylphosphonate"
XX
XX
XX MO200029596-A1.
XX
XX 25-MAY-2000.
XX
XX 16-NOV-1999; 99MO-US27152.
XX
XX 17-NOV-1998; 98US-0108763.
XX
XX (MONS) MONSANTO CO.
XX
XX Barry GF;
XX
XX MPI; 2000-387806/33.
XX
XX P-PSDB; AAT71250.
XX
XX Enhancing phosphonate herbicide tolerance in corn, tobacco, wheat,
PT cotton, canola and rice plants involves transforming plants with
PT phosphonate metabolizing genes encoding acyltransferase enzyme -
XX
XX Disclosure; Page 143-150; 179pp; English.
XX
XX The patent discloses a method for selectively enhancing the phosphonate
CC herbicide tolerance in recombinant corn, tobacco, wheat, cotton, canola
CC and rice plants, by transforming the plants with an expression cassette.
CC It comprises of a structural DNA sequence, that encodes an aminomethyl
CC phosphonic acid (AMPA) acyltransferase or transacylase (AAT), capable of
CC N-acylation of AMPA. This sequence has an amino terminal chloroplast
CC transit peptide (CTP), that targets AAT to the chloroplast. Co-expression
CC of glyphosate oxidase (GOX) gene, encoding glyphosate oxidoreductase,
CC along with AAT provides the transformed plants with higher resistance to
CC herbicides such as glyphosate and AMPA tolerance. In recombinant plants
CC and to prevent self-fertilisation and enhance hetero-fertilisation.
CC The present DNA sequence is the Escherichia coli phn operon sequence #2,
CC comprising the phnD gene, encoding AMPA acyltransferase enzyme (AAT).
XX
SQ Sequence 11672 BP; 2448 A; 3590 C; 3379 G; 2255 T; 0 other;

Query Match 2.9%; Score 17; DB 21; Length 11672;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 cctgttcagcgcagc 260
|||||
DB 4799 CCTGTTTCAGCGCCAGC 4783

RESULT 40
AAD01008/c
ID AAD01008 standard; DNA; 15611 BP.
XX
XX AAD01008;
XX
XX 21-SEP-2000 (first entry)
XX
XX Escherichia coli phn operon sequence #1.
XX
XX Phosphonate herbicide tolerance; aminomethyl phosphonic acid; AMPA; AAT;
KM acyltransferase; transacylase; recombinant plant; expression cassette;
KM corn; tobacco; wheat; cotton; canola; rice; chloroplast transit peptide;
KM CTP; glyphosate oxidase; GOX; glyphosate oxidoreductase; phn operon;
KM self-fertilisation; hetero-fertilisation; ds.
XX
XX Escherichia coli.
XX
OS
XX
FH Key Location/Qualifiers
FT CDS 13962..14396
FT /tag- a


```

FT      /*tag- ag
FT      /label="T"
FT      primer_bind
FT      39556..39574
FT      /*tag- ah
FT      primer_bind
FT      39877..39896
FT      /*tag- ai
FT      primer_bind
FT      39925..39943
FT      /*tag- aj
FT      primer_bind
FT      39932..39956
FT      /*tag- ak
FT      /note- "binds probe"
FT      replace(39944,T)
FT      allele
FT      /*tag- al
FT      primer_bind
FT      complement(39945..39963)
FT      /*tag- am
FT      primer_bind
FT      39953..39970
FT      /*tag- an
FT      primer_bind
FT      39954..39972
FT      /*tag- ao
FT      primer_bind
FT      39961..39985
FT      /*tag- ap
FT      /note- "binds probe"
FT      replace(39973,C)
FT      allele
FT      /*tag- aq
FT      primer_bind
FT      complement(39974..39992)
FT      /*tag- ar
FT      primer_bind
FT      40242..40259
FT      /*tag- as
FT      primer_bind
FT      41137..41154
FT      /*tag- at
FT      primer_bind
FT      41366..41384
FT      /*tag- au
FT      primer_bind
FT      41373..71397
FT      /*tag- av
FT      /note- "binds probe"
FT      replace(41385,C)
FT      allele
FT      /*tag- aw
FT      primer_bind
FT      41385..41403
FT      /*tag- ax
FT      primer_bind
FT      complement(41386..41404)
FT      /*tag- ay
FT      primer_bind
FT      41392..41416
FT      /*tag- az
FT      /note- "binds probe"
FT      replace(41404,C)
FT      allele
FT      /*tag- ba
FT      primer_bind
FT      complement(41405..41423)
FT      /*tag- bb
FT      primer_bind
FT      41564..41581
FT      /*tag- bc
FT      primer_bind
FT      42122..42141
FT      /*tag- bd
FT      primer_bind
FT      42213..42231
FT      /*tag- be
FT      primer_bind
FT      42220..42244
FT      /*tag- bf
FT      /note- "binds probe"
FT      replace(42232,C)
FT      allele
FT      /*tag- bg
FT      primer_bind
FT      complement(42233..42251)
FT      /*tag- bh
FT      primer_bind
FT      42526..42543
FT      /*tag- bi
FT      exon
FT      50436..50545
FT      /*tag- bj
FT      /label="G"
FT      primer_bind
FT      67289..67309
FT      /*tag- bk
FT      primer_bind
FT      67456..67474
FT      /*tag- bl
FT      primer_bind
FT      67463..67487
FT      /*tag- bm
FT      /note- "binds probe"

```

```

FT      allele
FT      replace(67475,G)
FT      /*tag- bn
FT      primer_bind
FT      complement(67476..67494)
FT      /*tag- bo
FT      primer_bind
FT      67724..67741
FT      /*tag- bp
FT      primer_bind
FT      69182..69200
FT      /*tag- bq
FT      primer_bind
FT      69502..69520
FT      /*tag- br
FT      primer_bind
FT      69509..69533
FT      /*tag- bs
FT      /note- "binds probe"
FT      allele
FT      /*tag- bt
FT      primer_bind
FT      replace(69521,G)
FT      /*tag- bu
FT      primer_bind
FT      complement(69522..69540)
FT      /*tag- bv
FT      primer_bind
FT      69609..69626
FT      /*tag- bv
FT      primer_bind
FT      72698..72715
FT      /*tag- bw
FT      primer_bind
FT      72819..72837
FT      /*tag- bx
FT      primer_bind
FT      72826..72850
FT      /*tag- by
FT      /note- "binds probe"
FT      allele
FT      /*tag- bz
FT      primer_bind
FT      complement(72839..72857)
FT      /*tag- ca
FT      primer_bind
FT      72881..72918
FT      /*tag- cb
FT      exon
FT      73099..73117
FT      /label="H"
FT      primer_bind
FT      /*tag- cc

```

```

Query Match      2.98; Score 17; DB 22; Length 240825;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      254 cgcagccaagcagcc 270
Db      23840 cgcagccaagcagcc 23856

```

```

RESULT 42
AA036842/c
ID AA036842 standard; DNA: 41 BP.
XX
AC AA036842;
XX
XX 16-JUN-1993 (first entry)
XX
DT 16-JUN-1993 (first entry)
XX
DE Adaptor-primer.
XX
KW Cyclophilin C; bone marrow; stromal cell line; AC 6; ss.
XX
OS Synthetic.
XX
PN WO9303050-A.
XX
PD 18-FEB-1993.
XX
PE 05-AUG-1992; 92WO-US06462.
XX
PR 05-AUG-1991; 91US-0740375.
XX
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
XX
PI Friedman JS, Weissman IL;
XX
DR WPI; 1993-076431/09.
XX

```

PT Cyclophilin C polypeptide and nucleic acid encoding it - useful
PT for screening a tissue-specific immunosuppressive agent

PS Example 1; Page 30; 63pp; English.

CC In the cDNA cloning and library construction described in Example
CC 1, total RNA was prep'd. from murine bone marrow derived stromal cell
CC line AC 6. cDNA synthesis was performed using an adaptor-primer
CC method (Rubenstein J. et al., (1990) Nuc. Acids Res. 18, 4833-4842)
CC with slight modifications. The adaptor-primer used, having the
CC sequence below, was the kind gift of Dr. Dan Denney.

XX
SQ Sequence 41 BP; 2 A; 9 C; 11 G; 19 T; 0 other;

Query Match 2.7%; Score 16; DB 14; Length 41;

Best Local Similarity 100.0%; Pred. No. 1.4e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 554 aaagagcctcgcgcgc 569
|||||
DB 27 AAAGGCTCCCGCGC 12

RESULT 43

AA073710 ID AA073710 standard; DNA; 222 BP.

XX AA073710;

DT 18-JUL-1995 (first entry)

XX Retrotransposon probe Tos1-1, cloned from Oryza sativa.

XX Retrotransposon; probe; rice; plant; identification; genetic uniformity;

KM pure line; Inbred parent line; near-isogenic line; ss.

XX Oryza sativa.

OS
FH Key Location/Qualifiers

FT Misc_feature 17..161

FT /*tag= a

FT /note= "probe sequence"

XX JP06277063-A.

XX 04-OCT-1994.

XX 08-OCT-1992; 92JP-0309167.

XX 08-OCT-1992; 92JP-0309167.

XX 08-OCT-1992; 92JP-0309167.

XX (NORI-) NORIN SUISANSO NOGYO SEIBUTSU SHIGEN.

PA (TAKE) TAKEDA CHEM IND LTD.

DR WPI; 1994-353746/44.

XX Identification of plants at the DNA level - using retrotransposon

PT sequences as probes

XX Claim 6; Page 5; 41pp; Japanese.

XX The nucleotide sequence of the retrotransposon probe Tos1-1 cloned from

CC the rice plant Oryza sativa. The probes may be used for the analysis of

CC plant genes to secure the efficient identification of plants and genetic

CC uniformity for the production of pure lines, Inbred parent lines and

CC near-isogenic lines.

XX Sequence 222 BP; 54 A; 48 C; 63 G; 57 T; 0 other;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 321 catgtcatctagtc 336
|||||

DB 34 catgtcatctagtc 49

RESULT 44

AA025170 ID AA025170 standard; DNA; 334 BP.

XX AA025170;

DT 19-JUL-1999 (first entry)

XX HIV-1 group O isolate 533 gp41 antigen DNA.

DE HIV-1 group O; Outlier strain; gp41; envelope protein; antigen;

KM vaccine; diagnosis; AIDS; ss.

XX Human immunodeficiency virus type 1.

OS WO9904011-A2.

XX 28-JAN-1999.

XX 20-JUL-1998; 98MO-EP04522.

XX 18-JUL-1997; 97EP-0870110.

XX (INNO-) INNOGENETICS NV.

XX Delaporte E, Peeters M, Saman E, Vanden Haesevelde M;

DR WPI; 1999-132255/11.

DR P-PSDB; AA05562.

XX New isolated HIV-1 group O strains - used to produce

PT polynucleotides, antigens and antibodies for use in diagnosis and in

PT vaccines for prevention of HIV-1 infection

XX Claim 4; Fig 6; 162pp; English.

XX This DNA sequence codes for an antigen (see AA05562) of the gp41

CC protein of HIV-1 group O (Outlier) strain 533, an isolate from

CC Gabon. The invention relates to new HIV-1 group O antigens (see

CC AA05562-625), and the use of these antigens, or nucleic acids

CC encoding them (see AA05562-80), in the diagnosis and prophylaxis of

CC AIDS. They can be used as reagents for detecting HIV-1 group O

CC infection and for differentiating different types of HIV-1 group O

CC infection. Vaccines that provide protective immunity against

CC HIV-1 infection, in particular against HIV-1 group O infection,

CC comprise at least one HIV-1 type O antigen, a nucleic acid encoding

CC such an antigen, a virus-like particle comprising such an antigen,

CC or an attenuated form of an HIV-1 type O strain. The invention also

CC relates to new HIV-1 group O strains, mostly from patients from

CC Cameroon and its neighbouring countries.

XX Sequence 334 BP; 124 A; 64 C; 77 G; 69 T; 0 other;

Query Match 2.7%; Score 16; DB 20; Length 334;

Best Local Similarity 100.0%; Pred. No. 1.3e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 540 accaacaagagaagaa 555
|||||

DB 290 accaacaagagaagaa 305

RESULT 45

AA021399/C ID AA021399 standard; DNA; 371 BP.

```

XX AAX21399;
AC
XX
XX 21-MAY-1999 (first entry)
XX
XX 3' fragment of AbsA operon from S.coelicolor.
DE
XX
XX AbsA; operon: open reading frame; histidine kinase; response regulator;
XX mutation; insertion; antibiotic; actinorhodin; undecylprodigiosin; ss.
XX
XX Streptomyces coelicolor.
OS
XX
XX Key Location/Qualifiers
XX CDS 1..371
XX /*tag= a
XX /note= "3' end of D9 coding sequence: no start codon
XX given at 5' end of sequence"
XX
XX US5876987-A.
XX
XX 02-MAR-1999.
XX
XX 06-FEB-1997; 96US-0796414.
XX
XX 07-FEB-1996; 96US-0011286.
XX 06-FEB-1997; 97US-0796414.
XX
XX (UNMS ) UNIV MICHIGAN STATE.
XX
XX PI Anderson TB, Brian P, Champness WC;
XX
XX WPI; 1999-189640/16.
XX
XX P-PSDB; AAW78494.
XX
XX
XX Recombinant Streptomyces species - useful for the production of
XX recombinant proteins and the hyper-expression of antibiotics,
XX particularly actinorhodin and undecylprodigiosin
XX
XX Claim 10; Fig 4; 30pp; English.
XX
XX The sequences AAX21397-X21399 represent overlapping sequences from the
XX AbsA operon of Streptomyces coelicolor. This sequence corresponds to
XX the downstream (3') fragment and contains a fragment of the D9 open
XX reading frame on the complementary strand to the Abs1 and Abs2 genes.
XX The Abs1 gene encodes a histidine kinase and the Abs2 gene encodes a
XX response regulator for the Abs1 gene. Inactivation of the Abs1 gene
XX either by mutation or insertion of heterologous DNA into the Abs1 gene
XX results in a Streptomyces organism which can hyperproduce heterologous
XX proteins or antibiotics, especially actinorhodin and undecylprodigiosin.
XX
XX Sequence 371 BP; 55 A; 141 C; 130 G; 45 T; 0 other;
XX
XX
XX Query Match 2.7%; Score 16; DB 20; Length 371;
XX Best Local Similarity 100.0%; Pred. No. 1.3e+02;
XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 506 ccagcgcgccgcgc 521
XX |||||||
XX Db 202 CCAAGCGCGCCTCCTC 187
XX
XX
XX RESULT 46
XX AAC01286/c
XX ID AAC01286 standard; cDNA; 398 BP.
XX
XX AAC01286;
XX
XX AC
XX
XX 06-OCT-2000 (first entry)
XX
XX Human secreted protein 5' EST, SEQ ID NO: 1284.
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX

```

```

KW gene therapy; chromosome mapping; ss.
XX
XX OS Homo sapiens.
XX
XX PN EP1033401-A2.
XX
XX PD 06-SEP-2000.
XX
XX PF 21-FEB-2000; 2000EP-0200610.
XX
XX PR 26-FEB-1999; 99US-0122487.
XX
XX (GEST ) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX
XX P-PSDB; AAG01280.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 1284; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. An ORF has been identified within the
XX sequence. The 5' ESTs were prepared from total human RNAs or POLYA+ RNAs
XX derived from 30 different tissues. EST sequences usually correspond
XX mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX well suited for isolating cDNA sequences derived from the 5' ends of
XX mRNAs and even in those cases where longer cDNA sequences have been
XX obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
XX cDNAs with intact 5' ends and can therefore be used to obtain full length
XX cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
XX gene therapy and chromosome mapping procedures. They are used to obtain
XX upstream regulatory sequences and to design expression and secretion
XX vectors.
XX
XX Sequence 398 BP; 86 A; 118 C; 110 G; 84 T; 0 other;
XX
XX
XX Query Match 2.7%; Score 16; DB 21; Length 398;
XX Best Local Similarity 100.0%; Pred. No. 1.3e+02;
XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 378 cactgagcgccgcgc 393
XX |||||||
XX Db 207 CACTGAGCGCGCGCC 192
XX
XX
XX RESULT 47
XX AAA79394/c
XX ID AAA79394 standard; cDNA; 413 BP.
XX
XX AAA79394;
XX
XX AC
XX
XX 27-NOV-2000 (first entry)
XX
XX EucaIyptus grandis cell signalling involved polynucleotide SEQ ID NO:195.
XX
XX DE EucaIyptus grandis; Pinus radiata; Monterey pine; plant; modification;
XX plant cell signalling; modulation; transgenic plant; pathogen; growth;
XX environmental change; development; cell proliferation; differentiation;
XX elongation; survival; disease resistance; nutrient metabolism; ss.
XX
XX OS EucaIyptus grandis.
XX
XX PN WO200042171-A1.
XX
XX 20-JUL-2000.
XX

```


PF 11-JAN-2000; 2000MO-US00724.
 XX
 PR 12-JAN-1999; 99US-0228986.
 PR 01-NOV-1999; 99US-0162866.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 PI Strabala TJ, Nieuwenhuizen NJ;
 DR WPI: 2000-476052/41.
 XX
 PT Isolated polynucleotide encoding a polypeptide involved in cell
 PT signaling used for generating transgenic plants with modified responses
 PT to external signals -
 XX
 PS Claim 1: Page 124; 527pp; English.
 XX
 CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
 CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
 CC pine (Pinus radiata also known as Monterey pine). The protein sequences
 CC are involved in cell signalling. The polynucleotide and protein
 CC sequences can be used to modify the response of plant cells to external
 CC signals e.g. environmental changes or pathogens during the growth and
 CC development of a plant. They can be used to modify cell proliferation,
 CC differentiation, elongation and survival, resistance to disease and
 CC nutrient metabolism. Examples of modifications which can be produced are
 CC altered fruit ripening and senescence of leaves and flowers e.g. to
 CC delay senescence and prolong the life of cut flowers or enhance
 CC senescence of reproductive organs to engineer sterile plants. Other
 CC modifications can be used to delay senescence in selected cell types or
 CC organs providing fruit and vegetables which have a longer shelf life
 CC between harvest and consumption, or to decrease branching frequency in
 CC forest tree species giving long stretches of valuable knot-free clear
 CC wood which can be used in solid timber furniture and veneers.
 XX
 SQ Sequence 413 BP; 101 A; 86 C; 118 G; 108 T; 0 other;

Query Match 2.7%; Score 16; DB 21; Length 413;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 acagcagccagcagca 129
 ||||||||||||||||
 DB 99 ACAGCAGCCAGGCCAA 84

RESULT 48
 AAC01726
 ID AAC01726 standard; cDNA; 489 BP.
 XX
 AC AAC01726;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein 5' EST, SEQ ID NO: 1724.
 XX
 KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KM gene therapy; chromosome mapping; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GENSET) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX
 DR WPI: 2000-500381/45.
 DR P-PSDB; AAC01720.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 1: SEQ ID 1724; 71pp + CD-ROM; English.
 XX
 CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. An ORF has been identified within the
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
 CC derived from 30 different tissues. EST sequences usually correspond
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 CC well suited for isolating cDNA sequences derived from the 5' ends of
 CC mRNAs and even in those cases where longer cDNA sequences have been
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 CC gene therapy and chromosome mapping procedures. They are used to obtain
 CC upstream regulatory sequences and to design expression and secretion
 CC vectors.
 XX
 SQ Sequence 489 BP; 146 A; 104 C; 146 G; 91 T; 2 other;

Query Match 2.7%; Score 16; DB 21; Length 489;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 542 cagcaggaagaaga 557
 ||||||||||||||||
 DB 351 caacaggaagaaga 366

RESULT 49
 AAT76907/C
 ID AAT76907 standard; DNA; 541 BP.
 XX
 AC AAT76907;
 XX
 DT 02-JUL-1998 (first entry)
 XX
 DE S. glaucescens Hstre DNA fragment.
 XX
 KM Acarbose biosynthesis: acbA gene; acbB gene; acbC gene; acbD gene;
 KM acbE gene; acbF gene; enzyme; alpha-amylase inhibitor; treatment;
 KM diabetes; Hstre gene; ss.
 XX
 OS Streptomyces glaucescens.
 XX
 PN DE19622783-A1.
 XX
 PD 11-DEC-1997.
 XX
 PF 07-JUN-1996; 96DE-1022783.
 XX
 PR 07-JUN-1996; 96DE-1022783.
 XX
 PA (FARH) HOECHST AG.
 XX
 PI Decker H;
 XX
 DR WPI: 1998-033827/04.
 XX
 PT Recombinant DNA molecule comprising genes for biosynthesis of
 PT acarbose - an alpha-amylase inhibitor useful in treatment of
 PT diabetes
 XX
 PS Example 2; Page 8; 35pp; German.
 XX

CC This sequence encodes a fragment of the *Streptomyces glaucescens* GLA.O
CC *hstre* gene which has been generated via PCR. This gene is used in a
CC method to identify the *acbd* gene which is involved in the acetate
CC biosynthesis pathway. *acbd* and other acetate biosynthesis genes such
CC as those represented in AAT75903 are useful for producing acetate.
CC which is an alpha-amylase inhibitor useful in the treatment of
CC diabetes.
XX
SQ Sequence 541 BP; 72 A; 173 C; 221 G; 75 T; 0 other;

Query Match 2.7%; Score 16; DB 19; Length 541;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 395 CGACGAGTCTACGCG 410
|||
DB 187 CGACGAGTCTACGCG 172

RESULT 50
AAT84039
ID AAT84039 standard; DNA: 624 BP.
XX
AC AAT84039;
XX
DT 01-SEP-1998 (first entry)
XX
DE DNA encoding a *Staphylococcus aureus* protein of unknown function.
XX
KM *Staphylococcus aureus* protein; ribozyme; antisense sequence; control;
KM *Staphylococcal* gene; regulatory element; bacterial gene expression;
KM vaccine; *Staphylococcal* infection; food poisoning; scaled skin syndrome;
KM toxic shock syndrome; ss.
XX
OS *Staphylococcus aureus*.
XX
FH Key Location/Qualifiers
FT CDS 187..477
FT /*tag= a
XX
PN WC9730070-A1.
XX
PD 21-AUG-1997.
XX
PF 19-FEB-1997; 97WO-US02318.
XX
PR 20-FEB-1996; 96US-0011888.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO;
PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;
XX
DR WPI: 1997-424969/39.
DR P-PSDB; AAM28105.
XX
PT Novel polypeptide(s) from *Staphylococcus aureus* strain WC9H29 - used
PT to isolate antimicrobial compounds, and in vaccines against *S.*
PT *aureus* infection
XX
PS Claim 9; Page 837; 989pp; English.
XX
XX The present sequence encodes a *Staphylococcus aureus* protein of
CC unknown function. The present sequence was isolated from a
CC library of clones of *S. aureus* WC9H 29 in *Escherichia coli*. The DNA
CC sequence can be used in the construction of ribozymes and antisense
CC sequences to control the expression of *Staphylococcal* genes. The DNA
CC sequence is also useful as a source of regulatory elements for the
CC control of bacterial gene expression. The encoded protein may be used
CC to produce vaccines to enable a host to produce specific antibodies
CC with antibacterial action. These vaccines and antibodies would protect
CC a host against invasion by *S. aureus*, and conditions relating to

CC *Staphylococcal* infection, e.g. *Staphylococcal* food poisoning, scaled
CC skin syndrome, and toxic shock syndrome.
XX

SQ Sequence 624 BP; 248 A; 72 C; 123 G; 171 T; 10 other;

Query Match 2.7%; Score 16; DB 18; Length 624;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 gtagtgagcgagcg 153
|||
DB 53 gtagtgagcgagcg 68

Search completed: September 21, 2001, 16:24:34
Job time: 9444 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 00:44:44 ; Search time 1872.43 Seconds
(without alignments)
2998.774 Million cell updates/sec

Title: US-09-138-735-1_COPY_1232_1825

Perfect score: 594
Sequence: 1 caggtacagcgtacgctt.....aaacgagctcgtgcgata 594

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 10228115 seqs, 4726426750 residues

Word size: 0

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*

190: gb_est110:*
191: gb_est111:*
192: gb_hlc:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_inv4:*
207: em_gss_inv5:*
208: em_gss_inv6:*
209: em_gss_inv7:*
210: em_gss_inv8:*
211: em_gss_inv9:*
212: em_gss_inv10:*
213: em_gss_inv11:*
214: em_gss_inv12:*
215: em_gss_inv13:*
216: em_gss_inv14:*
217: em_gss_inv15:*
218: em_gss_inv16:*
219: em_gss_inv17:*
220: em_gss_inv18:*
221: em_gss_inv19:*
222: em_gss_inv20:*
223: em_gss_inv21:*
224: em_gss_inv22:*
225: em_gss_inv23:*
226: em_gss_inv24:*
227: em_gss_inv25:*
228: em_gss_inv26:*
229: em_gss_inv27:*
230: em_gss_inv28:*
231: em_gss_inv29:*
232: em_gss_inv30:*
233: em_gss_inv31:*
234: em_gss_inv32:*
235: em_gss_inv33:*
236: em_gss_inv34:*
237: em_gss_inv35:*
238: em_gss_inv36:*
239: em_gss_inv37:*
240: em_gss_inv38:*
241: em_gss_inv39:*
242: em_gss_inv40:*
243: em_gss_inv41:*
244: em_gss_inv42:*
245: em_gss_inv43:*
246: em_gss_inv44:*
247: em_gss_inv45:*
248: em_gss_inv46:*
249: em_gss_inv47:*
250: em_gss_inv48:*
251: em_gss_inv49:*
252: em_gss_inv50:*
253: em_gss_inv51:*
254: em_gss_inv52:*
255: em_gss_inv53:*
256: em_gss_inv54:*
257: em_gss_inv55:*
258: em_gss_inv56:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|----------|---------------------|---------------------|
| 1 | 523 | 227 | AQ37527 | HS_5017_A | AQ37527 HS_5017_A |
| 2 | 656 | 245 | A249787 | LM0334G24 | A249787 LM0334G24 |
| 3 | 657 | 235 | AQ938059 | NR6-908R | AQ938059 NR6-908R |
| 4 | 713 | 221 | CNS04812 | AL278720 Tetradon | AL278720 Tetradon |
| 5 | 747 | 18 | AI325737 | AI325737 mm95d03.y | AI325737 mm95d03.y |
| 6 | 891 | 221 | CNS0330E | AL266639 Tetradon | AL266639 Tetradon |
| 7 | 235 | 124 | BB043716 | BB043716 BB043716 | BB043716 BB043716 |
| 8 | 261 | 6 | AA358477 | AA358477 EST67411 | AA358477 EST67411 |
| 9 | 327 | 122 | AM896020 | AM896020 QV4-NN003 | AM896020 QV4-NN003 |
| 10 | 334 | 104 | AI991901 | AI991901 w42b06.x | AI991901 w42b06.x |
| 11 | 334 | 104 | AM682518 | AM682518 EST01367 | AM682518 EST01367 |
| 12 | 392 | 157 | D78804 | D78804 HUM516H08B | D78804 HUM516H08B |
| 13 | 392 | 157 | AQ253544 | AQ253544 HS_2227_A | AQ253544 HS_2227_A |
| 14 | 463 | 104 | AI991951 | AI991951 w43a04.x | AI991951 w43a04.x |
| 15 | 463 | 151 | BF612300 | BF612300 dea17a11 | BF612300 dea17a11 |
| 16 | 465 | 140 | BE811867 | BE811867 PMO-AN003 | BE811867 PMO-AN003 |
| 17 | 474 | 244 | A2462258 | A2462258 LM0269023 | A2462258 LM0269023 |
| 18 | 476 | 20 | AI466074 | AI466074 vw39b02.y | AI466074 vw39b02.y |
| 19 | 493 | 164 | BE215806 | BE215806 HY_CEB000 | BE215806 HY_CEB000 |
| 20 | 498 | 136 | BE487307 | BE487307 175980 BA | BE487307 175980 BA |
| 21 | 512 | 164 | BE215810 | BE215810 HY_CEB000 | BE215810 HY_CEB000 |
| 22 | 527 | 228 | AQ399589 | AQ399589 m9xb0018E | AQ399589 m9xb0018E |
| 23 | 527 | 112 | AW168137 | AW168137 x96d04.x | AW168137 x96d04.x |
| 24 | 536 | 112 | AI158964 | AI158964 u142905.x | AI158964 u142905.x |
| 25 | 556 | 174 | BG147743 | BG147743 mab53h05. | BG147743 mab53h05. |
| 26 | 570 | 108 | AU164787 | AU164787 AI164787 | AU164787 AI164787 |
| 27 | 598 | 151 | BF620011 | BF620011 HYSMEC000 | BF620011 HYSMEC000 |
| 28 | 618 | 107 | AU101340 | AU101340 AU101340 | AU101340 AU101340 |
| 29 | 682 | 165 | BE291467 | BE291467 601085211 | BE291467 601085211 |
| 30 | 707 | 152 | BG323623 | BG323623 602422094 | BG323623 602422094 |
| 31 | 711 | 247 | AZ631126 | AZ631126 LM0485G09 | AZ631126 LM0485G09 |
| 32 | 712 | 251 | AZ867529 | AZ867529 2M0178H23 | AZ867529 2M0178H23 |
| 33 | 722 | 141 | BE892353 | BE892353 601433812 | BE892353 601433812 |
| 34 | 862 | 146 | BF27443 | BF27443 HY_CRA001 | BF27443 HY_CRA001 |
| 35 | 898 | 220 | AQ901727 | AQ901727 HS_2019_A | AQ901727 HS_2019_A |
| 36 | 909 | 220 | CNS01XEX | AL171604 Tetradon | AL171604 Tetradon |
| 37 | 919 | 221 | CNS030E7 | AL255866 Tetradon | AL255866 Tetradon |
| 38 | 952 | 222 | CNS05JPS | AL340489 Tetradon | AL340489 Tetradon |
| 39 | 968 | 139 | BE741337 | BE741337 601594184 | BE741337 601594184 |
| 40 | 986 | 220 | CNS035RQ | AL229103 Tetradon | AL229103 Tetradon |
| 41 | 1009 | 221 | CNS0305J | AL255520 Tetradon | AL255520 Tetradon |
| 42 | 1056 | 222 | CNS05EXA | AL333451 Tetradon | AL333451 Tetradon |
| 43 | 1078 | 106 | AL543701 | AL543701 AL543701 | AL543701 AL543701 |
| 44 | 1081 | 222 | CNS054X1 | AL321310 Tetradon | AL321310 Tetradon |
| 45 | 1101 | 192 | AK006925 | AK006925 Mus. muscu | AK006925 Mus. muscu |
| 46 | 1464 | 175 | BF118779 | BF118779 601755194 | BF118779 601755194 |
| 47 | 1481 | 175 | BG250472 | BG250472 602362637 | BG250472 602362637 |
| 48 | 164 | 23 | AI701300 | AI701300 wc59h07.x | AI701300 wc59h07.x |
| 49 | 165 | 1 | AA002343 | AA002343 mj06a07.r | AA002343 mj06a07.r |
| 50 | 206 | 1 | AA042040 | AA042040 m942a03.x | AA042040 m942a03.x |
| 51 | 207 | 28 | AV278443 | AV278443 AV278443 | AV278443 AV278443 |
| 52 | 212 | 122 | AM896035 | AM896035 QV4-NN003 | AM896035 QV4-NN003 |
| 53 | 219 | 238 | AZ083101 | AZ083101 RPT-23-4 | AZ083101 RPT-23-4 |
| 54 | 226 | 8 | AA473731 | AA473731 v989f03.r | AA473731 v989f03.r |
| 55 | 240 | 26 | AV196116 | AV196116 AV196116 | AV196116 AV196116 |
| 56 | 241 | 119 | AM647187 | AM647187 EST325730 | AM647187 EST325730 |
| 57 | 241 | 155 | C05473 | C05473 C05473 Humu | C05473 C05473 Humu |
| 58 | 244 | 144 | BF116745 | BF116745 uy91f04.y | BF116745 uy91f04.y |
| 59 | 248 | 189 | T84122 | T84122 yd33e12.r1 | T84122 yd33e12.r1 |
| 60 | 251 | 7 | AA416280 | AA416280 vd18b10.s | AA416280 vd18b10.s |
| 61 | 251 | 187 | N87803 | N87803 K0130F Humu | N87803 K0130F Humu |
| 62 | 256 | 227 | AV260749 | AV260749 AV260749 | AV260749 AV260749 |
| 63 | 260 | 141 | BE860963 | BE860963 UT-M-AN1- | BE860963 UT-M-AN1- |
| 64 | 264 | 6 | AA382428 | AA382428 EST95633 | AA382428 EST95633 |
| 65 | 265 | 161 | BB563790 | BB563790 BB563790 | BB563790 BB563790 |
| 66 | 266 | 6 | AA383032 | AA383032 EST96333 | AA383032 EST96333 |
| 67 | 266 | 162 | BB604316 | BB604316 BB604316 | BB604316 BB604316 |

| | | | | | | | | | | | | | |
|-------|----|-----|-----|-----|-----------|------------|-------|----|-----|-----|-----|-----------|-----------------------|
| C 141 | 17 | 2.9 | 420 | 110 | AM004870 | w029g12.x | C 214 | 17 | 2.9 | 500 | 236 | AO987028 | RPCT-23-2 |
| C 142 | 17 | 2.9 | 420 | 137 | BE588131 | WHE0657_G | C 215 | 17 | 2.9 | 502 | 22 | AI561151 | AI561151 tq37a10.x |
| C 143 | 17 | 2.9 | 421 | 155 | BC605585 | WHE2233_F | C 216 | 17 | 2.9 | 502 | 103 | AI928180 | AI928180 w095b01.x |
| C 144 | 17 | 2.9 | 421 | 168 | BF721024 | ma6b1h03 | C 217 | 17 | 2.9 | 503 | 240 | A2220281 | A2220281 Shared D |
| C 145 | 17 | 2.9 | 421 | 238 | AK084555 | RPCT-23-2 | C 218 | 17 | 2.9 | 503 | 241 | A2299040 | A2299040 RPCT-23-1 |
| C 146 | 17 | 2.9 | 423 | 12 | AA838588 | oe91a12.s | C 219 | 17 | 2.9 | 507 | 7 | AA433618 | AA433618 v4f47c03.r |
| C 147 | 17 | 2.9 | 424 | 139 | BE728941 | 601562379 | C 220 | 17 | 2.9 | 507 | 110 | AV756064 | AV756064 uj61h02.y |
| C 148 | 17 | 2.9 | 425 | 9 | AA589956 | vm14f07.r | C 221 | 17 | 2.9 | 509 | 18 | AI316056 | AI316056 uj61h02.y |
| C 149 | 17 | 2.9 | 425 | 16 | AI121590 | uc43f06.r | C 222 | 17 | 2.9 | 510 | 4 | AA221041 | AA221041 mw62h12.r |
| C 150 | 17 | 2.9 | 429 | 117 | AM563122 | 660072c11 | C 223 | 17 | 2.9 | 513 | 159 | N56635 | N56635 yw73h03.r |
| C 151 | 17 | 2.9 | 433 | 12 | AA799815 | EST189312 | C 224 | 17 | 2.9 | 518 | 145 | BF146275 | BF146275 TCAG31 Tr |
| C 152 | 17 | 2.9 | 433 | 32 | AV665130 | AV665130 | C 225 | 17 | 2.9 | 519 | 145 | AI346483 | AI346483 qp51a02.x |
| C 153 | 17 | 2.9 | 437 | 102 | AI808618 | wf56f05.x | C 226 | 17 | 2.9 | 519 | 104 | AI397440 | AI397440 AJ397440 |
| C 154 | 17 | 2.9 | 437 | 143 | BF011586 | us35a07.y | C 227 | 17 | 2.9 | 521 | 106 | AL587744 | AL587744 AL587744 |
| C 155 | 17 | 2.9 | 438 | 8 | AA477040 | aa26e12.r | C 228 | 17 | 2.9 | 522 | 238 | AA2110340 | RPCT-23-1 |
| C 156 | 17 | 2.9 | 439 | 19 | AI348257 | qp58d09.x | C 229 | 17 | 2.9 | 524 | 119 | AM654937 | MA AM654937 105443 MA |
| C 157 | 17 | 2.9 | 440 | 18 | AI265786 | AI265786 | C 230 | 17 | 2.9 | 524 | 138 | BE665288 | BE665288 153467 MA |
| C 158 | 17 | 2.9 | 441 | 7 | AA423608 | ve77b04.r | C 231 | 17 | 2.9 | 527 | 170 | BF826015 | BF826015 MR2-HN003 |
| C 159 | 17 | 2.9 | 441 | 246 | AA586342 | 1M0392011 | C 232 | 17 | 2.9 | 530 | 154 | BE502333 | BE502333 602550216 |
| C 160 | 17 | 2.9 | 442 | 117 | AM502469 | UI-HF-BR0 | C 233 | 17 | 2.9 | 531 | 189 | W07735 | W07735 zb02e10.r1 |
| C 161 | 17 | 2.9 | 442 | 190 | W24818 | zb05d08.r1 | C 234 | 17 | 2.9 | 532 | 19 | AI369027 | AI369027 qw27f05.x |
| C 162 | 17 | 2.9 | 444 | 1 | AA033458 | ml38a04.r | C 235 | 17 | 2.9 | 538 | 165 | BE238825 | BE238825 946009E12 |
| C 163 | 17 | 2.9 | 444 | 189 | TP80744 | TP80744 | C 236 | 17 | 2.9 | 539 | 119 | AM720808 | AM720808 833007E02 |
| C 164 | 17 | 2.9 | 445 | 110 | AM014376 | UI-H-B10- | C 237 | 17 | 2.9 | 541 | 32 | AV664061 | AV664061 AV664061 |
| C 165 | 17 | 2.9 | 445 | 158 | H91880 | y881c03.r1 | C 238 | 17 | 2.9 | 546 | 155 | BE557930 | BE557930 EM1_39_LDO |
| C 166 | 17 | 2.9 | 447 | 158 | BE666683 | 150283 MA | C 239 | 17 | 2.9 | 548 | 122 | AM968454 | AM968454 EST380530 |
| C 167 | 17 | 2.9 | 448 | 138 | BE759515 | an_2338 A | C 240 | 17 | 2.9 | 556 | 32 | AV665131 | AV665131 AV665131 |
| C 168 | 17 | 2.9 | 453 | 139 | BE137524 | BE137524 | C 241 | 17 | 2.9 | 557 | 7 | AA456966 | AA456966 aa90603.s |
| C 169 | 17 | 2.9 | 453 | 163 | BE137524 | BE137524 | C 242 | 17 | 2.9 | 559 | 173 | BC084490 | BC084490 H3101E06- |
| C 170 | 17 | 2.9 | 456 | 19 | AI341973 | q98a04.x | C 243 | 17 | 2.9 | 562 | 12 | AA833332 | AA833332 ud05h03.r |
| C 171 | 17 | 2.9 | 457 | 20 | AI1446668 | tj11f04.x | C 244 | 17 | 2.9 | 563 | 138 | BE665781 | BE665781 154938 MA |
| C 172 | 17 | 2.9 | 458 | 1 | AA002916 | mg38e10.r | C 245 | 17 | 2.9 | 567 | 190 | W72164 | W72164 zc47f110.r1 |
| C 173 | 17 | 2.9 | 459 | 17 | AI195909 | ue51b11.r | C 246 | 17 | 2.9 | 571 | 120 | AM743969 | AM743969 uj62b04.y |
| C 174 | 17 | 2.9 | 459 | 190 | W20518 | zb06g02.r1 | C 247 | 17 | 2.9 | 574 | 9 | AA619825 | AA619825 v161d05.s |
| C 175 | 17 | 2.9 | 460 | 241 | AA295346 | RPCT-23-1 | C 248 | 17 | 2.9 | 575 | 137 | BE583063 | BE583063 7-SE-WX P |
| C 176 | 17 | 2.9 | 461 | 10 | AA638263 | v197d03.r | C 249 | 17 | 2.9 | 576 | 115 | AA424454 | AA424454 707014P12 |
| C 177 | 17 | 2.9 | 461 | 23 | AI651742 | wb26c04.x | C 250 | 17 | 2.9 | 577 | 190 | W73840 | W73840 zc450d01.r1 |
| C 178 | 17 | 2.9 | 464 | 135 | BE555936 | d123f06.x | C 251 | 17 | 2.9 | 577 | 242 | A2357838 | A2357838 1M0099M08 |
| C 179 | 17 | 2.9 | 465 | 106 | AL554955 | AL554955 | C 252 | 17 | 2.9 | 579 | 245 | A2557695 | A2557695 RPCT-23-1 |
| C 180 | 17 | 2.9 | 465 | 187 | R23666 | y932a09.r1 | C 253 | 17 | 2.9 | 580 | 105 | AL506266 | AL506266 AL506266 |
| C 181 | 17 | 2.9 | 467 | 4 | AA260407 | va95f05.r | C 254 | 17 | 2.9 | 586 | 111 | AM076317 | AM076317 614065E04 |
| C 182 | 17 | 2.9 | 468 | 143 | BF042315 | BP250011B | C 255 | 17 | 2.9 | 588 | 166 | BE372802 | BE372802 601224318 |
| C 183 | 17 | 2.9 | 468 | 143 | BF042315 | BP250011B | C 256 | 17 | 2.9 | 589 | 111 | AM106080 | AM106080 um22h02.y |
| C 184 | 17 | 2.9 | 469 | 152 | BC370615 | BC370615 | C 257 | 17 | 2.9 | 589 | 166 | BE332979 | BE332979 us96h01.y |
| C 185 | 17 | 2.9 | 470 | 15 | AI021158 | ud02c10.r | C 258 | 17 | 2.9 | 593 | 156 | C95922 | C95922 C95922 Marc |
| C 186 | 17 | 2.9 | 470 | 112 | AM140903 | EST290837 | C 259 | 17 | 2.9 | 593 | 155 | BE542875 | BE542875 602571050 |
| C 187 | 17 | 2.9 | 471 | 21 | AI549442 | AI549442 | C 260 | 17 | 2.9 | 593 | 165 | C95923 | C95923 C95923 Marc |
| C 188 | 17 | 2.9 | 472 | 20 | AI447169 | AI447169 | C 261 | 17 | 2.9 | 593 | 165 | BE237905 | BE237905 894038H08 |
| C 189 | 17 | 2.9 | 474 | 136 | BE494727 | BE494727 | C 262 | 17 | 2.9 | 593 | 167 | BE448856 | BE448856 ut88f09.y |
| C 190 | 17 | 2.9 | 475 | 21 | AI553844 | AI553844 | C 263 | 17 | 2.9 | 593 | 242 | A2374514 | A2374514 1M01227M17 |
| C 191 | 17 | 2.9 | 476 | 22 | AI580492 | AI580492 | C 264 | 17 | 2.9 | 597 | 32 | AV708855 | AV708855 AV708855 |
| C 192 | 17 | 2.9 | 477 | 102 | AI818992 | wj60f01.x | C 265 | 17 | 2.9 | 603 | 117 | AM546477 | AM546477 L0008G06- |
| C 193 | 17 | 2.9 | 479 | 144 | BF130957 | 601819779 | C 266 | 17 | 2.9 | 606 | 226 | AO257474 | AO257474 nbdb0018E |
| C 194 | 17 | 2.9 | 479 | 188 | R99845 | R99845 | C 267 | 17 | 2.9 | 612 | 165 | BE291736 | BE291736 6011084456 |
| C 195 | 17 | 2.9 | 481 | 9 | AA571593 | vm07b12.r | C 268 | 17 | 2.9 | 617 | 14 | AA968297 | AA968297 uc71g08.y |
| C 196 | 17 | 2.9 | 481 | 12 | AA819761 | AA819761 | C 269 | 17 | 2.9 | 617 | 122 | AM912093 | AM912093 u4f1f06.y |
| C 197 | 17 | 2.9 | 481 | 236 | AO9777390 | AO9777390 | C 270 | 17 | 2.9 | 618 | 8 | AA343452 | AA343452 vj82a05.r |
| C 198 | 17 | 2.9 | 482 | 6 | AA397099 | mx84b12.r | C 271 | 17 | 2.9 | 618 | 32 | AAV17648 | AAV17648 AV717648 |
| C 199 | 17 | 2.9 | 482 | 158 | H30789 | yo79f07.r1 | C 272 | 17 | 2.9 | 619 | 154 | BE472976 | BE472976 602514920 |
| C 200 | 17 | 2.9 | 483 | 17 | AI230840 | EST227535 | C 273 | 17 | 2.9 | 621 | 137 | BE566092 | BE566092 6025139091 |
| C 201 | 17 | 2.9 | 483 | 165 | BE301059 | BE301059 | C 274 | 17 | 2.9 | 625 | 18 | AI316963 | AI316963 u124g10.y |
| C 202 | 17 | 2.9 | 487 | 148 | BR444477 | BR444477 | C 275 | 17 | 2.9 | 625 | 152 | BC354973 | BC354973 00685_16a |
| C 203 | 17 | 2.9 | 487 | 154 | BE516219 | EL5TEd45 | C 276 | 17 | 2.9 | 626 | 142 | BE960901 | BE960901 601648138 |
| C 204 | 17 | 2.9 | 488 | 142 | BE987729 | UI-M-CG0P | C 277 | 17 | 2.9 | 636 | 256 | B54011 | B54011 C1T-HSP-201 |
| C 205 | 17 | 2.9 | 490 | 175 | BC239824 | OVI_29_E0 | C 278 | 17 | 2.9 | 637 | 122 | AM898205 | AM898205 RC3-NN006 |
| C 206 | 17 | 2.9 | 491 | 122 | AM952892 | EST364962 | C 279 | 17 | 2.9 | 638 | 152 | BC370097 | BC370097 LMA105 V8 |
| C 207 | 17 | 2.9 | 491 | 144 | BF122166 | BF122166 | C 280 | 17 | 2.9 | 640 | 138 | BE648606 | BE648606 UI-M-BH2 |
| C 208 | 17 | 2.9 | 492 | 245 | A2521734 | RPCT-11-1 | C 281 | 17 | 2.9 | 640 | 243 | A2407409 | A2407409 1M0178005 |
| C 209 | 17 | 2.9 | 495 | 13 | AA879755 | vv98c07.r | C 282 | 17 | 2.9 | 644 | 242 | A2356482 | A2356482 1M0097006 |
| C 210 | 17 | 2.9 | 496 | 7 | AA467505 | ve01c09.r | C 283 | 17 | 2.9 | 646 | 247 | A2646241 | A2646241 1M0512006 |
| C 211 | 17 | 2.9 | 497 | 111 | AM106034 | um22c05.y | C 284 | 17 | 2.9 | 648 | 1 | AA044836 | AA044836 zk74d10.r |
| C 212 | 17 | 2.9 | 497 | 138 | BE624612 | BE624612 | C 285 | 17 | 2.9 | 648 | 241 | A2332758 | A2332758 1M0061P22 |
| C 213 | 17 | 2.9 | 500 | 150 | BE544886 | UI-R-A0-a | C 286 | 17 | 2.9 | 652 | 236 | A2005606 | A2005606 RPCT-23-2 |

| | | | | | | | | | | | | | |
|-------|----|-----|-----|-----|----------|-----------|-----|----|-----|-----|-----|----------|----------|
| C 433 | 16 | 2.7 | 282 | 114 | AM322270 | uo55f09.x | 506 | 16 | 2.7 | 373 | 108 | AV030288 | AV030288 |
| C 434 | 16 | 2.7 | 282 | 162 | BB603068 | BB603068 | 507 | 16 | 2.7 | 374 | 16 | AI153832 | AI153832 |
| C 435 | 16 | 2.7 | 283 | 190 | W63249 | W63249 | 508 | 16 | 2.7 | 376 | 147 | BF366887 | BF366887 |
| C 436 | 16 | 2.7 | 284 | 172 | BG014672 | BG014672 | 509 | 16 | 2.7 | 376 | 235 | AO933991 | AO933991 |
| C 437 | 16 | 2.7 | 285 | 5 | AA319807 | AA319807 | 510 | 16 | 2.7 | 377 | 115 | AA041253 | AA041253 |
| C 438 | 16 | 2.7 | 285 | 114 | AM293702 | AM293702 | 511 | 16 | 2.7 | 379 | 2 | AA134777 | AA134777 |
| C 439 | 16 | 2.7 | 285 | 172 | BG014680 | BG014680 | 512 | 16 | 2.7 | 381 | 223 | AO027089 | AO027089 |
| C 440 | 16 | 2.7 | 287 | 246 | AA625139 | AA625139 | 513 | 16 | 2.7 | 381 | 223 | AO028370 | AO028370 |
| C 441 | 16 | 2.7 | 288 | 128 | BB228806 | BB228806 | 514 | 16 | 2.7 | 382 | 3 | AA201950 | AA201950 |
| C 442 | 16 | 2.7 | 292 | 113 | AM230121 | AM230121 | 515 | 16 | 2.7 | 382 | 144 | BF116581 | BF116581 |
| C 443 | 16 | 2.7 | 292 | 134 | BB452180 | BB452180 | 516 | 16 | 2.7 | 383 | 116 | AA430242 | AA430242 |
| C 444 | 16 | 2.7 | 293 | 8 | AA539651 | AA539651 | 517 | 16 | 2.7 | 383 | 226 | AO292706 | AO292706 |
| C 445 | 16 | 2.7 | 294 | 162 | BB607769 | BB607769 | 518 | 16 | 2.7 | 386 | 2 | AA105799 | AA105799 |
| C 446 | 16 | 2.7 | 300 | 26 | AV184952 | AV184952 | 519 | 16 | 2.7 | 386 | 16 | AI158599 | AI158599 |
| C 447 | 16 | 2.7 | 300 | 132 | BB384543 | BB384543 | 520 | 16 | 2.7 | 386 | 105 | AL510254 | AL510254 |
| C 448 | 16 | 2.7 | 301 | 120 | AM802255 | AM802255 | 521 | 16 | 2.7 | 387 | 251 | AZ927302 | AZ927302 |
| C 449 | 16 | 2.7 | 302 | 6 | AA365537 | AA365537 | 522 | 16 | 2.7 | 388 | 112 | AM177000 | AM177000 |
| C 450 | 16 | 2.7 | 302 | 21 | AI549953 | AI549953 | 523 | 16 | 2.7 | 389 | 139 | BE724104 | BE724104 |
| C 451 | 16 | 2.7 | 303 | 190 | W64779 | W64779 | 524 | 16 | 2.7 | 390 | 1 | AA034582 | AA034582 |
| C 452 | 16 | 2.7 | 305 | 167 | BE38830 | BE38830 | 525 | 16 | 2.7 | 390 | 114 | AM282576 | AM282576 |
| C 453 | 16 | 2.7 | 307 | 192 | AK006677 | AK006677 | 526 | 16 | 2.7 | 391 | 235 | AO872539 | AO872539 |
| C 454 | 16 | 2.7 | 313 | 12 | AA822603 | AA822603 | 527 | 16 | 2.7 | 391 | 251 | AZ880172 | AZ880172 |
| C 455 | 16 | 2.7 | 313 | 109 | AV045222 | AV045222 | 528 | 16 | 2.7 | 395 | 10 | AA680351 | AA680351 |
| C 456 | 16 | 2.7 | 314 | 108 | AV022870 | AV022870 | 529 | 16 | 2.7 | 396 | 172 | BF988683 | BF988683 |
| C 457 | 16 | 2.7 | 314 | 191 | Z24498 | Z24498 | 530 | 16 | 2.7 | 396 | 237 | AZ048323 | AZ048323 |
| C 458 | 16 | 2.7 | 315 | 238 | AA118498 | AA118498 | 531 | 16 | 2.7 | 397 | 18 | AI298998 | AI298998 |
| C 459 | 16 | 2.7 | 317 | 11 | AA783378 | AA783378 | 532 | 16 | 2.7 | 397 | 7 | AA454208 | AA454208 |
| C 460 | 16 | 2.7 | 320 | 25 | AV106214 | AV106214 | 533 | 16 | 2.7 | 399 | 16 | AI116992 | AI116992 |
| C 461 | 16 | 2.7 | 320 | 102 | AI806593 | AI806593 | 534 | 16 | 2.7 | 399 | 170 | BF833779 | BF833779 |
| C 462 | 16 | 2.7 | 324 | 250 | AA857898 | AA857898 | 535 | 16 | 2.7 | 400 | 224 | AO143095 | AO143095 |
| C 463 | 16 | 2.7 | 326 | 258 | P422L | P422L | 536 | 16 | 2.7 | 401 | 146 | BE286220 | BE286220 |
| C 464 | 16 | 2.7 | 329 | 24 | AI770014 | AI770014 | 537 | 16 | 2.7 | 401 | 146 | BE286220 | BE286220 |
| C 465 | 16 | 2.7 | 329 | 138 | BE672691 | BE672691 | 538 | 16 | 2.7 | 405 | 13 | AA921685 | AA921685 |
| C 466 | 16 | 2.7 | 330 | 110 | AV746311 | AV746311 | 539 | 16 | 2.7 | 406 | 14 | AI209184 | AI209184 |
| C 467 | 16 | 2.7 | 331 | 14 | AA946593 | AA946593 | 540 | 16 | 2.7 | 406 | 17 | AA196706 | AA196706 |
| C 468 | 16 | 2.7 | 331 | 158 | H53514 | H53514 | 541 | 16 | 2.7 | 407 | 12 | AA816828 | AA816828 |
| C 469 | 16 | 2.7 | 333 | 158 | H90383 | H90383 | 542 | 16 | 2.7 | 407 | 107 | AO091598 | AO091598 |
| C 470 | 16 | 2.7 | 335 | 1 | AA008392 | AA008392 | 543 | 16 | 2.7 | 407 | 107 | BE499769 | BE499769 |
| C 471 | 16 | 2.7 | 335 | 23 | AI645058 | AI645058 | 544 | 16 | 2.7 | 409 | 241 | AZ302507 | AZ302507 |
| C 472 | 16 | 2.7 | 335 | 168 | BE709340 | BE709340 | 545 | 16 | 2.7 | 410 | 220 | CNS02DP3 | CNS02DP3 |
| C 473 | 16 | 2.7 | 335 | 190 | W33553 | W33553 | 546 | 16 | 2.7 | 416 | 17 | AI237778 | AI237778 |
| C 474 | 16 | 2.7 | 336 | 15 | AI065368 | AI065368 | 547 | 16 | 2.7 | 417 | 4 | AA238136 | AA238136 |
| C 475 | 16 | 2.7 | 336 | 114 | AM291945 | AM291945 | 548 | 16 | 2.7 | 417 | 30 | AV397717 | AV397717 |
| C 476 | 16 | 2.7 | 339 | 13 | AA938349 | AA938349 | 549 | 16 | 2.7 | 413 | 167 | BE439110 | BE439110 |
| C 477 | 16 | 2.7 | 339 | 141 | BE860869 | BE860869 | 550 | 16 | 2.7 | 416 | 21 | AI552688 | AI552688 |
| C 478 | 16 | 2.7 | 340 | 105 | AL511056 | AL511056 | 551 | 16 | 2.7 | 416 | 115 | AA419046 | AA419046 |
| C 479 | 16 | 2.7 | 340 | 146 | BE286341 | BE286341 | 552 | 16 | 2.7 | 417 | 7 | AA436989 | AA436989 |
| C 480 | 16 | 2.7 | 341 | 14 | AA965555 | AA965555 | 553 | 16 | 2.7 | 417 | 141 | BE847940 | BE847940 |
| C 481 | 16 | 2.7 | 342 | 170 | BR883586 | BR883586 | 554 | 16 | 2.7 | 417 | 21 | AI524631 | AI524631 |
| C 482 | 16 | 2.7 | 343 | 102 | AI847406 | AI847406 | 555 | 16 | 2.7 | 417 | 110 | AM035593 | AM035593 |
| C 483 | 16 | 2.7 | 343 | 258 | TA33E0SP | TA33E0SP | 556 | 16 | 2.7 | 418 | 115 | AA419046 | AA419046 |
| C 484 | 16 | 2.7 | 344 | 248 | AA713132 | AA713132 | 557 | 16 | 2.7 | 419 | 8 | AA505554 | AA505554 |
| C 485 | 16 | 2.7 | 345 | 12 | AA844266 | AA844266 | 558 | 16 | 2.7 | 419 | 14 | AA965710 | AA965710 |
| C 486 | 16 | 2.7 | 351 | 116 | AA480447 | AA480447 | 559 | 16 | 2.7 | 419 | 141 | BE847940 | BE847940 |
| C 487 | 16 | 2.7 | 353 | 158 | HA2425 | HA2425 | 560 | 16 | 2.7 | 419 | 226 | AO263146 | AO263146 |
| C 488 | 16 | 2.7 | 354 | 2 | AA096628 | AA096628 | 561 | 16 | 2.7 | 419 | 233 | AO805822 | AO805822 |
| C 489 | 16 | 2.7 | 355 | 113 | AM214689 | AM214689 | 562 | 16 | 2.7 | 420 | 30 | AV431288 | AV431288 |
| C 490 | 16 | 2.7 | 356 | 119 | AM663133 | AM663133 | 563 | 16 | 2.7 | 420 | 154 | BG516182 | BG516182 |
| C 491 | 16 | 2.7 | 358 | 140 | BE819977 | BE819977 | 564 | 16 | 2.7 | 423 | 234 | AO844136 | AO844136 |
| C 492 | 16 | 2.7 | 358 | 144 | BF071087 | BF071087 | 565 | 16 | 2.7 | 424 | 225 | AO229268 | AO229268 |
| C 493 | 16 | 2.7 | 359 | 233 | AO803883 | AO803883 | 566 | 16 | 2.7 | 425 | 123 | AA856459 | AA856459 |
| C 494 | 16 | 2.7 | 360 | 26 | AV197870 | AV197870 | 567 | 16 | 2.7 | 427 | 237 | AZ073851 | AZ073851 |
| C 495 | 16 | 2.7 | 360 | 156 | CA7720 | CA7720 | 568 | 16 | 2.7 | 427 | 244 | AZ489703 | AZ489703 |
| C 496 | 16 | 2.7 | 361 | 104 | BF934505 | BF934505 | 569 | 16 | 2.7 | 428 | 1 | AA051534 | AA051534 |
| C 497 | 16 | 2.7 | 361 | 174 | AI978332 | AI978332 | 570 | 16 | 2.7 | 428 | 14 | AA989827 | AA989827 |
| C 498 | 16 | 2.7 | 361 | 111 | AM100950 | AM100950 | 571 | 16 | 2.7 | 428 | 172 | BF990176 | BF990176 |
| C 499 | 16 | 2.7 | 362 | 7 | AA438431 | AA438431 | 572 | 16 | 2.7 | 429 | 9 | AA576188 | AA576188 |
| C 500 | 16 | 2.7 | 363 | 7 | AA407556 | AA407556 | 573 | 16 | 2.7 | 429 | 154 | BE508736 | BE508736 |
| C 501 | 16 | 2.7 | 363 | 114 | AM325063 | AM325063 | 574 | 16 | 2.7 | 430 | 258 | P971L | P971L |
| C 502 | 16 | 2.7 | 370 | 14 | AA969069 | AA969069 | 575 | 16 | 2.7 | 431 | 145 | BF201611 | BF201611 |
| C 503 | 16 | 2.7 | 370 | 189 | T99585 | T99585 | 576 | 16 | 2.7 | 432 | 117 | AI268354 | AI268354 |
| C 504 | 16 | 2.7 | 371 | 23 | AI697579 | AI697579 | 577 | 16 | 2.7 | 432 | 118 | AA558166 | AA558166 |
| C 505 | 16 | 2.7 | 371 | 138 | BE675255 | BE675255 | 578 | 16 | 2.7 | 432 | 225 | AO173762 | AO173762 |

| | | | | | | | | | | | | | |
|-------|----|-----|-----|-----|------------|--------------------|-------|----|-----|-----|-----|-----------|--------------------|
| C 725 | 16 | 2.7 | 523 | 251 | A2894737 | RPCT-24-2 | C 798 | 16 | 2.7 | 567 | 173 | BC076107 | BC076107 H315G04- |
| C 726 | 16 | 2.7 | 524 | 119 | AM656571 | 108893 MA | C 799 | 16 | 2.7 | 568 | 146 | BF302411 | BF302411 602031415 |
| C 727 | 16 | 2.7 | 524 | 233 | AO784935 | HS_3251.A | C 800 | 16 | 2.7 | 569 | 10 | AA689770 | AA689770 vs07F04.F |
| C 728 | 16 | 2.7 | 525 | 166 | BE375872 | 601229688 | C 801 | 16 | 2.7 | 570 | 173 | BC077002 | BC077002 H3009F11- |
| C 729 | 16 | 2.7 | 526 | 166 | BE301858 | BB01d05.Y | C 802 | 16 | 2.7 | 571 | 9 | AA628975 | AA628975 zU73a03.s |
| C 730 | 16 | 2.7 | 526 | 248 | A2733499 | RPCT-24-1 | C 803 | 16 | 2.7 | 571 | 165 | BE287588 | BE287588 601097348 |
| C 731 | 16 | 2.7 | 527 | 1 | AA039173 | m198e10.r | C 804 | 16 | 2.7 | 573 | 17 | A122039 | A122039 601097348 |
| C 732 | 16 | 2.7 | 527 | 15 | A1039144 | ox31b09.s | C 805 | 16 | 2.7 | 573 | 20 | A1436401 | A1436401 f10e12.x |
| C 733 | 16 | 2.7 | 527 | 15 | A1052714 | o226d09.x | C 806 | 16 | 2.7 | 573 | 112 | AM131852 | AM131852 xF35d11.x |
| C 734 | 16 | 2.7 | 528 | 139 | BE727230 | 601563139 | C 807 | 16 | 2.7 | 574 | 111 | AM105656 | AM105656 um29g04.Y |
| C 735 | 16 | 2.7 | 528 | 258 | TA270F030 | AL486931 T. brucei | C 808 | 16 | 2.7 | 577 | 113 | AM268594 | AM268594 xv34h06.x |
| C 736 | 16 | 2.7 | 529 | 14 | AA979135 | LD33409.5 | C 809 | 16 | 2.7 | 579 | 145 | BE200417 | BE200417 WHE2257.H |
| C 737 | 16 | 2.7 | 529 | 102 | A1790940 | UK29912.Y | C 810 | 16 | 2.7 | 580 | 147 | BE348830 | BE348830 QV4-DT002 |
| C 738 | 16 | 2.7 | 529 | 113 | AM231706 | AW331706 687018H09 | C 811 | 16 | 2.7 | 581 | 121 | AM811944 | AM811944 RC2-ST016 |
| C 739 | 16 | 2.7 | 530 | 106 | AT002867 | AT002867 | C 812 | 16 | 2.7 | 582 | 23 | A1679735 | A1679735 fu76b01.x |
| C 740 | 16 | 2.7 | 530 | 243 | A2436805 | 1M0224P08 | C 813 | 16 | 2.7 | 582 | 231 | AO654365 | AO654365 Sheared D |
| C 741 | 16 | 2.7 | 531 | 17 | A1245004 | q194b01.x | C 814 | 16 | 2.7 | 583 | 152 | BC375097 | BC375097 UI-R-CV1- |
| C 742 | 16 | 2.7 | 531 | 119 | AM679708 | WS1_30_LB0 | C 815 | 16 | 2.7 | 583 | 165 | BE293614 | BE293614 601186819 |
| C 743 | 16 | 2.7 | 531 | 173 | BC064432 | BC064432 H3019F02- | C 816 | 16 | 2.7 | 585 | 3 | AA202231 | AA202231 LD02412.5 |
| C 744 | 16 | 2.7 | 531 | 1 | AA057248 | z156f01.r | C 817 | 16 | 2.7 | 585 | 7 | AA439071 | AA439071 LD13520.5 |
| C 745 | 16 | 2.7 | 533 | 13 | AA874604 | AA874604 vx78b10.r | C 818 | 16 | 2.7 | 585 | 137 | BE598987 | BE598987 P11_86_A0 |
| C 746 | 16 | 2.7 | 533 | 246 | AZ586537 | 1M0392102 | C 819 | 16 | 2.7 | 588 | 102 | A1809035 | A1809035 wh74e04.x |
| C 747 | 16 | 2.7 | 534 | 138 | BE676666 | 7F34a05.x | C 820 | 16 | 2.7 | 588 | 166 | BE307600 | BE307600 601091435 |
| C 748 | 16 | 2.7 | 534 | 251 | A2918311 | BE945122 | C 821 | 16 | 2.7 | 588 | 168 | BF716322 | BF716322 saal8e04. |
| C 749 | 16 | 2.7 | 535 | 171 | BF945122 | PM1-NM120 | C 822 | 16 | 2.7 | 589 | 168 | AM25011 | AM25011 WS1_74.CO |
| C 750 | 16 | 2.7 | 536 | 110 | AM012285 | AM012285 | C 823 | 16 | 2.7 | 590 | 152 | BC377597 | BC377597 602435144 |
| C 751 | 16 | 2.7 | 536 | 166 | BE364670 | PI1_14.DO | C 824 | 16 | 2.7 | 591 | 166 | BE367025 | BE367025 P11_42_B0 |
| C 752 | 16 | 2.7 | 536 | 167 | BE426118 | WHE0328.E | C 825 | 16 | 2.7 | 592 | 121 | AM811867 | AM811867 RC2-ST016 |
| C 753 | 16 | 2.7 | 537 | 10 | AA631353 | AA631353 n92h11.s | C 826 | 16 | 2.7 | 592 | 227 | AO374864 | AO374864 PC111-15 |
| C 754 | 16 | 2.7 | 537 | 240 | AZ223372 | RPCT-23-7 | C 827 | 16 | 2.7 | 593 | 111 | AM105986 | AM105986 um20f09.Y |
| C 755 | 16 | 2.7 | 538 | 107 | AU128824 | AU128824 | C 828 | 16 | 2.7 | 593 | 113 | AM249214 | AM249214 2820819.s |
| C 756 | 16 | 2.7 | 538 | 174 | BG224979 | kp52g09.Y | C 829 | 16 | 2.7 | 594 | 222 | FR0041102 | FR0041102 |
| C 757 | 16 | 2.7 | 538 | 258 | P330L | AA539506 | C 830 | 16 | 2.7 | 595 | 256 | BE5051 | BE5051 C1R-HSP-201 |
| C 758 | 16 | 2.7 | 539 | 102 | A1791088 | A1791088 | C 831 | 16 | 2.7 | 595 | 256 | BE5051 | BE5051 C1R-HSP-201 |
| C 759 | 16 | 2.7 | 539 | 173 | BG058497 | uK55e11.Y | C 832 | 16 | 2.7 | 597 | 147 | BE941780 | BE941780 EST256354 |
| C 760 | 16 | 2.7 | 540 | 31 | AV608383 | AV608383 | C 833 | 16 | 2.7 | 597 | 165 | BE272728 | BE272728 601104917 |
| C 761 | 16 | 2.7 | 540 | 167 | BE418912 | BE418912 WMR017.AL | C 834 | 16 | 2.7 | 598 | 8 | AA544785 | AA544785 vx35e06.r |
| C 762 | 16 | 2.7 | 541 | 32 | AV668225 | AV668225 | C 835 | 16 | 2.7 | 599 | 120 | AM747621 | AM747621 WS1_74.CO |
| C 763 | 16 | 2.7 | 541 | 137 | BE544769 | 601075534 | C 836 | 16 | 2.7 | 599 | 222 | FR0023577 | FR0023577 |
| C 764 | 16 | 2.7 | 541 | 251 | A2918326 | A2918326 | C 837 | 16 | 2.7 | 600 | 6 | AA390672 | AA390672 LD09763.5 |
| C 765 | 16 | 2.7 | 542 | 9 | AA587084 | AA587084 n77h12.s | C 838 | 16 | 2.7 | 600 | 23 | A1684380 | A1684380 tc96g02.x |
| C 766 | 16 | 2.7 | 542 | 113 | AM202672 | f120e02.Y | C 839 | 16 | 2.7 | 600 | 115 | AM422817 | AM422817 f148d12.Y |
| C 767 | 16 | 2.7 | 543 | 7 | AA439857 | LD14536.5 | C 840 | 16 | 2.7 | 600 | 222 | FR0038454 | FR0038454 |
| C 768 | 16 | 2.7 | 543 | 111 | AM106558 | AM106558 | C 841 | 16 | 2.7 | 605 | 231 | AO651163 | AO651163 Sheared D |
| C 769 | 16 | 2.7 | 543 | 145 | BF151743 | BF151743 u219h04.Y | C 842 | 16 | 2.7 | 605 | 237 | AZ012653 | AZ012653 RPCT-23-3 |
| C 770 | 16 | 2.7 | 545 | 24 | A1767579 | u1767c11.x | C 843 | 16 | 2.7 | 606 | 111 | AM106098 | AM106098 um23b02.Y |
| C 771 | 16 | 2.7 | 546 | 155 | BG549982 | 947079H10 | C 844 | 16 | 2.7 | 606 | 152 | BC355152 | BC355152 947043D04 |
| C 772 | 16 | 2.7 | 547 | 30 | AV391046 | AV391046 | C 845 | 16 | 2.7 | 608 | 227 | AO383759 | AO383759 RPCT11-12 |
| C 773 | 16 | 2.7 | 547 | 149 | BE469236 | BE469236 | C 846 | 16 | 2.7 | 610 | 222 | FR0038454 | FR0038454 |
| C 774 | 16 | 2.7 | 547 | 162 | BE014769 | BE014769 | C 847 | 16 | 2.7 | 611 | 152 | BC351960 | BC351960 |
| C 775 | 16 | 2.7 | 548 | 258 | TA1313E04P | TA1313E04P | C 848 | 16 | 2.7 | 611 | 222 | FR0017452 | FR0017452 |
| C 776 | 16 | 2.7 | 549 | 20 | A1466676 | A1466676 m87bD03.Y | C 849 | 16 | 2.7 | 613 | 156 | C76640 | C76640 C76640 Mous |
| C 777 | 16 | 2.7 | 549 | 235 | AO892278 | AO892278 HS_2025.B | C 850 | 16 | 2.7 | 614 | 105 | AL502085 | AL502085 |
| C 778 | 16 | 2.7 | 550 | 116 | AM494336 | AM494336 UI-M-BH3- | C 851 | 16 | 2.7 | 614 | 166 | BE364374 | BE364374 |
| C 779 | 16 | 2.7 | 550 | 113 | AM257996 | AM257996 | C 852 | 16 | 2.7 | 615 | 166 | BE364374 | BE364374 |
| C 780 | 16 | 2.7 | 552 | 113 | AM257996 | AM257996 | C 853 | 16 | 2.7 | 615 | 111 | AM099789 | AM099789 sd31e07.Y |
| C 781 | 16 | 2.7 | 553 | 165 | BE290024 | 601089630 | C 854 | 16 | 2.7 | 615 | 173 | BC048518 | BC048518 CVL14_E0 |
| C 782 | 16 | 2.7 | 554 | 175 | AM542242 | AM542242 | C 855 | 16 | 2.7 | 616 | 113 | AM211243 | AM211243 u129e02.Y |
| C 783 | 16 | 2.7 | 554 | 175 | BC240707 | C0155C09- | C 856 | 16 | 2.7 | 616 | 167 | BE458850 | BE458850 EST141442 |
| C 784 | 16 | 2.7 | 555 | 149 | BF493066 | BF493066 | C 857 | 16 | 2.7 | 617 | 173 | BC052764 | BC052764 |
| C 785 | 16 | 2.7 | 556 | 123 | AM977277 | AM977277 | C 858 | 16 | 2.7 | 618 | 2 | AA137524 | AA137524 m931b01.r |
| C 786 | 16 | 2.7 | 556 | 244 | AZ4940308 | AZ4940308 | C 859 | 16 | 2.7 | 618 | 32 | AV714382 | AV714382 |
| C 787 | 16 | 2.7 | 558 | 110 | AM044866 | um14903.Y | C 860 | 16 | 2.7 | 618 | 122 | AM957874 | AM957874 |
| C 788 | 16 | 2.7 | 558 | 173 | BC082833 | BC082833 | C 861 | 16 | 2.7 | 619 | 13 | AA886661 | AA886661 o147f09.s |
| C 789 | 16 | 2.7 | 560 | 104 | A1956204 | u172f109.Y | C 862 | 16 | 2.7 | 619 | 245 | AZ507408 | AZ507408 |
| C 790 | 16 | 2.7 | 560 | 155 | BC610490 | 326432 MA | C 863 | 16 | 2.7 | 620 | 8 | AA541089 | AA541089 LD20881.5 |
| C 791 | 16 | 2.7 | 560 | 165 | BE240640 | EST044689 | C 864 | 16 | 2.7 | 620 | 112 | AM1774096 | AM1774096 |
| C 792 | 16 | 2.7 | 564 | 102 | BF492168 | BF492168 | C 865 | 16 | 2.7 | 621 | 149 | BF484185 | BF484185 |
| C 793 | 16 | 2.7 | 565 | 102 | A1790877 | UK29902.Y | C 866 | 16 | 2.7 | 623 | 114 | AA986912 | AA986912 uc08a07.Y |
| C 794 | 16 | 2.7 | 565 | 105 | AL120577 | DKF2p761D | C 867 | 16 | 2.7 | 623 | 114 | AM319327 | AM319327 un11e04.Y |
| C 795 | 16 | 2.7 | 565 | 240 | AZ216815 | Sheared D | C 868 | 16 | 2.7 | 624 | 223 | AO014308 | AO014308 RPCT11-22 |
| C 796 | 16 | 2.7 | 565 | 249 | AZ791826 | AZ791826 | C 869 | 16 | 2.7 | 625 | 110 | AV763030 | AV763030 |
| C 797 | 16 | 2.7 | 566 | 117 | AM544651 | AM544651 | C 870 | 16 | 2.7 | 625 | 147 | BF320596 | BF320596 u254h03.Y |


```

REFERENCE      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS      Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
              Keller,A., Shaker,R., Furlong,J.J., Young,J., Zhao,S., Adams,M.D., and
              Hood,L.
TITLE        Sequence-tagged connectors: A sequence approach to mapping and
              scanning the human genome
JOURNAL      Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE      99380589
COMMENT      Contact: Mahairas GG, Wallace JC, Hood L
              High Throughput Sequencing Center
              University of Washington
              401 Queen Anne Avenue North, Seattle, WA 98109, USA
              Tel: (206) 616-3618
              Fax: (206) 616-3887
              Email: jwallace@u.washington.edu
              Clones may be purchased from Research Genetics (info@resgen.com).
              BAC end Web Server: http://www.husc.washington.edu
              Plate: 593 Row: K Column: 4
              Seq primer: T7
              Class: BAC ends
              High quality sequence stop: 523.
FEATURES
  source
    1..523
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_plate="593 Col=4 Row=K"
    /clone_lib="RPC11 Human Male BAC Library"
    /sex="Male"
    /cell_type="Lymphocytes"
    /note="Vector: pBAC6.6; RPC11 Human Male BAC Library"
BASE COUNT   163 a 151 c 125 g 58 t 6 others
ORIGIN
Query Match      3.2%; Score 19; DB 227; Length 523;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 347 ggataatgctcaccacaag 365
|||||
Db 151 GGATTAAATGTCACCCAAG 169
|||||
RESULT 2
LOCUS      A2497878 656 bp DNA GSS 05-OCT-2000
DEFINITION 1M0334G248 Mouse 10kb plasmid U06C1M library Mus musculus genomic
clone U06C1M0334G24 R, DNA sequence.
ACCESSION  A2497878
VERSION    A2497878
KEYWORDS
SOURCE     GSS.
ORGANISM  Mus musculus.
          house mouse.
          Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 656)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

```

| FEATURES | source |
|--|--|
| Plate: 0334 | row: G column: 24 |
| Seq primer: | CACACAGGACAGCAGTATGACC |
| Class: | plasmid ends |
| High quality sequence stop: | 656. |
| Location/Qualifiers | 1. .656 |
| | /organism="Mus musculus" |
| | /strain="C57BL/6J" |
| | /db.xref="taxon:10090" |
| | /clone="U08100334G24" |
| | /clone.lib="Mouse 10kb plasmid library" |
| | /sex="male" |
| | /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-" |
| | /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson laboratory Mouse DNA Resource |
| | (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g147321419b1A7129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance." |
| BASE COUNT | 155 a 176 c 132 g 192 t 1 others |
| ORIGIN | |
| Query Match | 3.2%; Score 19; DB 245; Length 656; |
| Best Local Similarity | 100.0%; Pred. No. 21; |
| Matches | 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| Oy | 537 gaaccacagagaaga 555 |
| Db | 110 GAACCAACAGCAGAGAA 92 |
| RESULT | 3 |
| LOCUS | A0938059 |
| DEFINITION | N65-908R Human NotI clones Homo sapiens genomic, DNA sequence. |
| ACCESSION | A0938059 |
| VERSION | A0938059.1 GI:7214437 |
| KEYWORDS | GSS. |
| SOURCE | human. |
| ORGANISM | Homo sapiens |
| Eukaryota; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | |
| REFERENCE | 1 (bases 1 to 657) |
| AUTHORS | Zabarovskiy,E.R., Glazutlin,R., Podowski,R.M., Zabarovskaya,V.V., Xie,L., Muravenko,O.V., Kozhev,S., Petrenko,L., Skobdeva,N., Li,J., Protolopov,A., Kashuba,V., Ernberg,I., Winberg,G. and Wahlstedt,C. |
| TITLE | NotI clones in the analysis of the human genome |
| JOURNAL | Nucleic Acids Res. 28 (7), 1635-1639 (2000) |
| MEDLINE | 20175728 |
| COMMENT | Contact: Podowski RM Center for Genomics Research Karolinska Institute 17177 Stockholm, Sweden Tel: +46-8-728-6372 Fax: +46-8-337983 Email: Raf.Podowski@cr.ki.se Class: NotI site. |
| FEATURES | Location/Qualifiers |
| source | 1. .657 |

[illegible]

| COMMENT | FEATURES | BASE COUNT | ORIGIN |
|--|---|-----------------------------|--|
| <p>Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil</p> <p>Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br</p> <p>This sequence was derived from the FAPSP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?fl=613-QV4-NN0039-110) 500-204-003613-2000-05-1164-1)</p> <p>Seq primer: puc 18 forward High quality sequence start: 9 High quality sequence stop: 132. Location/Qualifiers 1..327</p> | <p>/organism="Homo sapiens" /db_xref="taxon:9606" /clone_id="NN0039" /dev_stage="Adult" /note="Organ: nervous.normal; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORSRES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."</p> | <p>107 a 65 c 98 g 57 t</p> | <p>Query Match Best Local Similarity 100.0%; Score 18; DB 122; Length 327; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;</p> |
| <p>RESULT 10 LOCUS A1991901/c DEFINITION ws42b06.x1 NCI-CGAP Brn25 Homo sapiens cDNA clone IMAGE:249823 3' similar to contains element MSRI repetitive element ;, mRNA sequence. ACCESSION A1991901 VERSION A1991901.1 GI:5838806 KEYWORDS EST. SOURCE human. ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 334) NCI/NIHDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BRGAP), Tumor Gene Index Unpublished (1998) Contact: Robert Strausberg, Ph.D. Email: cgapsr@mail.nih.gov Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonalato, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNLT at: www-bio.lnlt.gov/brp/image/image.html Insert Length: 555 Std Error: 0.00 Seq primer: -400P from Glibco</p> | | | |

MEDLINE
 99380589
COMMENT
 Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Sequence Tagged Connector
 Plate: 2227 row: M column: 8
 Class: BAC ends
 High quality sequence stop: 392.
FEATURES
 source
 1..392
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CIT Approved Human Genomic Sperm Library D"
 /sex="male"
 /note="Organ: sperm; Vector: pBelOABCl1; BAC clones in E-Coli DH10B"
BASE COUNT
 94 a 101 c 67 g 129 t 1 others
ORIGIN
 Query Match 3.0%; Score 18; DB 226; Length 392;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 489 ggcagctaatagtcgtgc 506
 ||||||||||||||||
Db 154 GGCAGCTAATAGTGTCTC 137
RESULT 14
 AI991951 463 bp mRNA EST 08-MAR-2000
 AI991951/c ws43a04.x1 NCI-CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2499918 3'
LOCUS
DEFINITION
 similar to contains element MSRI repetitive element ;, mRNA
 sequence.
ACCESSION
 AI991951 GI:5838856
VERSION
 AI991951
KEYWORDS
 EST.
SOURCE
 human.
ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
 1 (bases 1 to 463)
 NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
 National Cancer Institute / National Institute of Neurological
 Disorders and Stroke, Brain Tumor Genome Anatomy Project
 (CGAP/BRGAP), Tumor Gene Index
 Unpublished (1998)
JOURNAL
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
COMMENT
 Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
 Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/dbrrp/image/image.html
 Insert length: 566 Std Error: 0.00
 Seq primer: -400P from Gdbco.
FEATURES
 source
 1..463
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:2499918"
 /clone_lib="NCI-CGAP_Brn25"
 /tissue_type="anaplastic oligodendroglioma"

FEATURES
 source
 1..463
 Location/Qualifiers
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a
 modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer (5'
 TGTTCACCAATCGAAGTGGAGCGCCGCAATAGTGTGTGTGTGTGTGT
 T 3'); double-stranded cDNA was ligated to Eco RI
 adaptors (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of the modified pT73 vector.
 Library is normalized, and was constructed by Bento
 Soares and M. Fatima Bonaldo."
BASE COUNT
 99 a 102 c 132 g 129 t 1 others
ORIGIN
 Query Match 3.0%; Score 18; DB 104; Length 463;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 230 cccctgcacacccccctg 247
 ||||||||||||||||
Db 151 CCCCTGCATCACCCCTG 134
RESULT 15
 BF612300 463 bp mRNA EST 14-DEC-2000
LOCUS
 BF612300 daa17a11.y1 NICHD XGC Lul Xenopus laevis cDNA clone IMAGE:4056476
DEFINITION
 5' similar to SW:GSP.FORCA P56101 CYSTEINE STRING PROTEIN ;, mRNA
 sequence.
ACCESSION
 BF612300
VERSION
 BF612300.1 GI:11782798
KEYWORDS
 EST.
SOURCE
 <African clawed frog.
ORGANISM
 Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 Xenopodinae; Xenopus.
REFERENCE
 1 (bases 1 to 463)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
JOURNAL
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
COMMENT
 Tissue Procurement: Martha Redbert, Steven L. Klein, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
 Seq primer: -40RP from Gdbco
 High quality sequence stop: 439.
FEATURES
 source
 1..463
 Location/Qualifiers
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="IMAGE:4056476"
 /clone_lib="NICHD XGC Lul"
 /dev_stage="adult"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: oligo dT.
 Average insert size 2.6 kb. Constructed by Life
 Technologies. Note: This is a Xenopus gene Collection (XGC
) library."
BASE COUNT
 108 a 111 c 138 g 106 t
ORIGIN
 Query Match 3.0%; Score 18; DB 151; Length 463;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 453 cgggacacatgcccacc 470
 |||||||
 Db 30 CCGGCACAAATGGCCGACC 47

RESULT 16
 BE811867 465 bp mRNA EST 21-SEP-2000
 LOCUS PM0-AN0032-110700-001-a12 AN0032 Homo sapiens CDNA, mRNA sequence.
 DEFINITION BE811867
 VERSION BE811867.1 GI:10244101
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 465)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R.,
 Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 COMMENT 20202663
 CONTACT: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=6t2-PM0-AN0032-110
 700-001-a12&t3=2000-07-11&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 266.
 Location/Qualifiers
 1. 465
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="AN0032"
 /dev_stage="Adult"
 /note="Organ: amnion,normal; Vector: puc18; Site:1: SmaI;
 Site:2: SmaI: A mini-library was made by cloning products
 derived from ORESSES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 111 a 160 c 92 g 102 t
 ORIGIN

Query Match 3.0%; Score 18; DB 140; Length 465;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 cccctgcacacccctg 247
 |||||||
 Db 138 CCCCTGCATCACCCTCG 155

RESULT 17
 A2462258 474 bp DNA GSS 04-OCT-2000
 LOCUS IM0269023F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 DEFINITION clone UUGC1M0269023 F, DNA sequence.
 ACCESSION A2462258

VERSION A2462258.1 GI:10620383
 GSS.
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 474)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 ,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
 and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 TITILE Unpublished (2000)
 JOURNAL Contact: Robert B. Weiss
 COMMENT University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 1000 Std Error: 0.00
 Plate: 0269 row: 0 column: 23
 Seq primer: CGTTGTAAACGACGCGCAGT
 Class: plasmid ends
 High quality sequence stop: 474.
 Location/Qualifiers
 1. 474
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0269023"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PM042ny; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared for a derivative
 of pMD42 (g147321141g147129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."
 BASE COUNT 85 a 132 c 135 g 122 t
 ORIGIN

Query Match 3.0%; Score 18; DB 244; Length 474;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 gcagccgcgcctcctg 282
 |||||||
 Db 219 GCAGCCGCGCGCTCTGCC 236

RESULT 18
 A1466074 476 bp mRNA EST 09-MAR-1999
 LOCUS VW39b02.y1 Soares mammary gland NDMWG Mus musculus cDNA clone
 DEFINITION IMAGE:1246155 5' similar to gb:U17574 Mouse mRNA for

| | | |
|--------------------------|---|--|
| ACCESSION | AI66074 | procrasin-binding protein (MOUSE);, mRNA sequence. |
| VERSION | AI66074.1 | GI:4320104 |
| KEYWORDS | EST. | |
| SOURCE | house mouse. | |
| ORGANISM | Mus musculus | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | |
| REFERENCE | 1 (bases 1 to 476) | |
| AUTHORS | NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . | |
| TITLE | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index | |
| JOURNAL | Unpublished (1997) | |
| COMMENT | Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:659843 | |
| FEATURES | This read is a RESSEQUENCE of a previously sequenced mouse clone This read has been verified (found to hit its original self in the correct orientation) Seq primer: -40RP from Glibco High quality sequence stop: 317. Location/Qualifiers 1..476 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="IMAGE:1246155" /clone_lib="Soares_mammary_gland_NbMWG" /sex="male" /tissue_type="mammary gland" /dev_stage="4 weeks" /lab_host="DH10B" /note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dt) primer [5' TGTTACCAATCTGATGATGGAGCGCCGCGAATGTTTTTTTTTTTTTTTTTTTTTT T 3'] ; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Patima BonaIdo." | |
| BASE COUNT | 104 a 132 c 134 g 106 t | |
| ORIGIN | | |
| Query Match | 3.0%; | Score 18; DB 20; Length 476; |
| Best Local Similarity | 100.0%; | Pred. No. 77; |
| Matches 18; Conservative | 0; | Mismatches 0; Indels 0; Gaps 0; |
| QY 265 | gcagccgcgcctccctgcgc 282 | |
| | | |
| DB 380 | gcagccgcgcctccctgcgc 397 | |
| RESULT 19 | | |
| BE215806/c | | |
| LOCUS | BE215806 | 493 bp mRNA EST 09-MAR-2001 |
| DEFINITION | HV.CE50008116f Hordeum vulgare seedling green leaf EST library | |
| | HV.CE50008116f, mRNA sequence. | |
| ACCESSION | BE215806 | |
| VERSION | BE215806.1 | GI:8903418 |
| KEYWORDS | EST. | |
| SOURCE | Barley. | |
| ORGANISM | Hordeum vulgare | |
| | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae ; Triticeae; Hordeum. | |
| REFERENCE | 1 (bases 1 to 493) | |

| | | | |
|--------------------------|---|---|-------------|
| AUTHORS | | Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Fritch,D., Yu,Y., Anderson,H., Dale,J., Henry,D., Kennedle,S., Palmer,M., Rambo,I., Sasaki,C., Schwartzbeck,U., Simmons,J., Choi,D.W., Main,D. and Wood,J. | |
| TITLE | | Development of a genetically and physically anchored EST resource for barley genomics | |
| JOURNAL | | Unpublished (2000) | |
| COMMENT | | Contact: Wing RA Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 7288 Fax: 864 656 4293 Email: rwing@clemson.edu Seq primer: AATTACCTCCTCACTTAAGG High quality sequence stop: 486. | |
| FEATURES | | location/Qualifiers | |
| source | | 1. 493 | |
| | | /organism="Hordeum vulgare" /cultivar="C116151 (ML6)" /db_xref="taxon:4513" /clone="HV_CED0008116f" /clone_lib="Hordeum vulgare seedling green leaf EST library HVCN0005 (Erysiphe infected & control)" /tissue_type="seedling green leaf" /lab_host="SOLR" /note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI" | |
| BASE COUNT | | 55 a 213 c 151 g 74 t | |
| ORIGIN | | | |
| Query Match | | 3.0%, Score 18; DB 164; Length 493; | |
| Best Local Similarity | | 100.0%; Pred. No. 77; | |
| Matches 18; Conservative | | 0; Mismatches 0; Indels 0; Gaps 0; | |
| Oy | 385 ggcgcggccacgacgag 402 | | |
| Db | 73 GCGCGCGCCACGACGAG 56 | | |
| RESULT 20 | | | |
| LOCUS | BE487307 | 498 bp | mRNA |
| DEFINITION | 175980 BABRC 5BOV Bos taurus cDNA 5', mRNA sequence. | EST | 28-AUG-2000 |
| ACCESSION | BE487307 | | |
| VERSION | BE487307.1 | GI:9606840 | |
| KEYWORDS | EST. | | |
| SOURCE | cow. | | |
| ORGANISM | Bos taurus | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos. | | |
| REFERENCE | 1 (bases 1 to 498) | | |
| AUTHORS | Sonstegard,T.S., Capuco,A.V., Van Tassel,C.P., Ashwell,M.S. and Wells,K.D. | | |
| TITLE | Mapping of Expressed Sequence Tags from a normalized bovine mammary gland cDNA library | | |
| JOURNAL | Unpublished (2000) | | |
| COMMENT | Contact: Sonstegard TS USDA, ARS, Beltsville Agricultural Research Center Bldg. 200 Km 2A, Beltsville, MD 20705, USA Tel: 301 504 8416 Fax: 301 504 8414 Email: tads@psl.barc.usda.gov Single pass sequencing. Bases called and alt. trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options. PCR primers FORWARD: AGCAACAGCTATGACCAT BACKWARD: GTTTCACGTACGACG Plate: 123 row: K column: 2 Seq primer: ATTTAGGTGACACTATATAC. Location/Qualifiers | | |
| FEATURES | | | |

```

source
1. 498
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone.lib="BARC SBOV"
/tissue_type="pooled"
/lab_host="DH10B"
/Note="Vector: PCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled mRNA isolated from mammary
tissues at eight physiological, developmental, and disease
states."

BASE COUNT      87 a      181 c      125 g      105 t

ORIGIN

Query Match      3.0%; Score 18; DB 136; Length 498;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 514 gccgtctccacagcc 531
|||||
Db 27 GCCTCGTCTCCACAGCC 44

RESULT 21
BE215810/c 512 bp mRNA EST 09-MAR-2001
LOCUS HV_CEBD008120f Hordeum vulgare seedling green leaf EST library
DEFINITION HVCDNA0005 (Erysiphe infected & control) Hordeum vulgare cDNA clone
HV_CEBD008120f, mRNA sequence.
ACCESSION BE215810
VERSION BE215810.1 GI:8903422
KEYWORDS EST.
SOURCE Hordeum vulgare
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Hordeum.
1 (bases 1 to 512)
Wing,R., Close,T.J., Kleinbofs,A., Wise,R., Begum,D., Frisch,D., Yu
,T., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo
,T., Sasaki,C., Schwartzbeck,J., Simmons,J., Choi,D.W., Main,D. and
Wood,T.
Development of a genetically and physically anchored EST resource
for barley genomics
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: AATTAACCTCCTCAAGAGG
High quality sequence start: 2
High quality sequence stop: 233.
Location/Qualifiers
1. 512
/organism="Hordeum vulgare"
/cultivar="C116151 (Mia6)"
/db_xref="taxon:4513"
/clone="HV_CEBD008120f"
/clone.lib="Hordeum vulgare seedling green leaf EST
library HVCDNA0005 (Erysiphe infected & control)"
/tissue_type="seedling green leaf"
/lab_host="SOLR"
/Note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT      114 a      167 c      127 g      104 t

ORIGIN

Query Match      3.0%; Score 18; DB 164; Length 512;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 385 gccgcgcgcacagagc 402
|||||
Db 75 GCCCGCGCCACGACGAG 58

RESULT 22
AQ399589 512 bp DNA GSS 06-MAR-1999
LOCUS mxb0018E01f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
DEFINITION clone mxb0018E01f, DNA sequence.
ACCESSION AQ399589
VERSION AQ399589.1 GI:4370616
KEYWORDS GSS.
SOURCE Magnaporthe grisea.
ORGANISM Magnaporthe grisea
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes Incertae sedis; Magnaporthaceae; Magnaporthe.
1 (bases 1 to 512)
Yu,Y., Zhu,H., Boyd,C.A., Gaudette,B., Gayle,A., Kingsbury,R.,
Phillips,K., Sasinowski,M., Wing,R.A. and Dean,R.A.
A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
Unpublished (1998)
Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdeane@clemson.edu
Seq primer: TAATGACTCCTATAGCG
Class: BAC ends
High quality sequence stop: 354.
Location/Qualifiers
1. 512
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mxb0018E01f"
/clone.lib="CUGI Rice Blast BAC Library"
/tissue_type="protoplasts"
/lab_host="E. coli DH10B"
/Note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25x genome coverage. High density colony filters
are available upon request."

BASE COUNT      102 a      167 c      131 g      112 t

ORIGIN

Query Match      3.0%; Score 18; DB 228; Length 512;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 271 gccgcctcgcgcgcgcgc 288
|||||
Db 239 GCCGCTCTGCGCGCGC 222

RESULT 23
AM168137 527 bp mRNA EST 12-NOV-1999
LOCUS xg60d04.x1 NCL_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:2632711 3',
DEFINITION mRNA sequence.
ACCESSION AM168137

```

VERSION AM168137.1 GI:6399662
 KEYWORDS EST
 SOURCE human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 527)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bdrr/image/image.html

FEATURES
 source
 1.527
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2632711"
 /clone_id="NCI-CGAP_UF4"
 /tissue_type="serous papillary carcinoma, high grade, 2 pooled tumors"
 /lab_host="DH10B"
 /note="Organ: uterus; Vector: PCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.48 kb. Life Technologies catalog #: 11542-016"

BASE COUNT 151 a 152 c 162 g 61 t 1 others
 ORIGIN

Query Match 3.0%; Score 18; DB 112; Length 527;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 160 ggggtgtgtctctcgggc 177
 ||||||||||||||||
 Db 469 GGGGTGTGTCTCTCGGC 486

RESULT 24
 A1158964/c 536 bp mRNA EST 02-OCT-1998
 DEFINITION u142g05.x1 Sugano mouse embryo mewa Mus musculus cDNA clone
 IMAGE:1885112 3' similar to WP:131B12.1 C505599 ;, mRNA sequence.
 ACCESSION A1158964
 VERSION A1158964.1 GI:3692146
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 536)
 AUTHORS Marra, M., Hallier, J., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Treising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project

Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:969436

FEATURES
 source
 1.536
 /organism="Mus musculus"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1885112"
 /clone_id="Sugano mouse embryo mewa"
 /dev_stage="embryo, 14 dpc"
 /lab_host="DH10B"
 /note="Vector: pME18S-Fl3; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTTCTGTG], digested and cloned into distinct DraIII sites of the pME18S-Fl3 vector (5' site CACCATGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCGCTCTAAAGCTGCG and 3' end primer CGACCTGCAGCTCAGCACA."

BASE COUNT 138 a 120 c 99 g 179 t
 ORIGIN

Query Match 3.0%; Score 18; DB 16; Length 536;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 543 aacagagaaagaaagc 560
 ||||||||||||||||
 Db 516 AACAGAGAAAGAAAGC 499

RESULT 25
 BG147743 556 bp mRNA EST 01-FEB-2001
 BG147743
 DEFINITION mab53h05.x1 Soares.NMBA_brachial_arch Mus musculus cDNA clone
 IMAGE:3974241 3' similar to TR:075462 075462 CYTOKINE-LIKE FACTOR-1 PRECURSOR. [1] ;, mRNA sequence.
 ACCESSION BG147743
 VERSION BG147743.1 GI:12651151
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 556)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

image.llnl.gov/image/html/iresources.shtml

MG1:1474273

Possible reversed clone: similarity on wrong strand

Possible reversed clone: polyT not found

High quality sequence stop: 463.

FEATURES

source

Location/Qualifiers

1..556
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3974241"
/clone_lib="Soares NMEBA branchial arch"
/tissue_type="branchial arches"
/dev_stage="embryo, 10.5 dpc"
/lab_host="DH10B (phage resistant)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: NotI; Site 2: EcoRI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTCACATCTGAGGTGGAGCGCGCCGATCATTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo."

BASE COUNT 101 a 183 c 162 g 110 t

ORIGIN

Query Match

Best Local Similarity 3.0%; Score 18; DB 174; Length 556;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 159 ggggtgtgtgtcctcg 176

|||||

Db 32 GGGGTGTGTCTCTCGG 49

RESULT 26

LOCUS

AU164787 570 bp mRNA EST 20-NOV-2000

DEFINITION

AU164787 Rice cDNA from immature leaf including apical meristem

(under short day condition) Oryza sativa cDNA clone E61272, mRNA

sequence.

ACCESSION

AU164787.1 GI:11233375

VERSION

KEYWORDS

SOURCE

ORGANISM

Oryza sativa.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Eurharoidae; Oryzaceae; Oryza.

1 (bases 1 to 570)

Sasaki, T. and Yamamoto, K.

Unpublished (2000)

Rice cDNA from immature leaf including apical meristem

Contact: Takuji Sasaki

National Institute of Agrobiological Resources

Rice Genome Research Program

2-1-2 Kannondai, Tsukuba

Ibaraki,

Japan 305

Tel: 0298-38-7441

Fax: 0298-38-7468

Email: tsasaki@abr.affrc.go.jp

PROJECT "RGP"

E61272_22.

FEATURES

source

Location/Qualifiers

1..570
/organism="Oryza sativa"
/strain="Nipponbare"
/db_xref="taxon:4530"
/clone="E61272"
/clone_lib="Rice cDNA from immature leaf including apical

meristem (under short day condition)"

/dev_stage="Immature leaf including apical meristem (under

short day condition)"

BASE COUNT 123 a 186 c 132 g 126 t 3 others

ORIGIN

Query Match

Best Local Similarity 3.0%; Score 18; DB 108; Length 570;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 394 acgacgaggtctacggcg 411

|||||

Db 243 ACGACGAGGTCTACGGCG 260

RESULT 27

LOCUS

BF620011 598 bp mRNA EST 22-FEB-2001

DEFINITION

HVSMC000924f Hordeum vulgare seedling shoot EST library

HVCNDA0003 (Etiolated and unstressed) Hordeum vulgare cDNA clone

HVSMC000924f, mRNA sequence.

ACCESSION

BF620011.2 GI:13107977

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Wing, R., Close, T.J., Kleinbols, A., Wise, R., Begum, D., Frisch, D., Yu

, Y., Anderson, H., Dale, J., Henry, D., Kernodle, S., Palmer, M., Rambo

, T., Sasaki, C., Schwartzbeck, J., Simmons, J., Choi, D.W., Main, D. and

Wood, T.

Development of a genetically and physically anchored EST resource

for barley genomics

Unpublished (2000)

On Dec 18, 2000 this sequence version replaced gi:11883745.

Contact: Wing RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Seq primer: AATTACCTCCTAAGCG

High quality sequence stop: 592.

FEATURES

source

Location/Qualifiers

1..598
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMC000924f"
/clone_lib="Hordeum vulgare seedling shoot EST library
HVCNDA0003 (Etiolated and unstressed)"
/tissue_type="Seedling shoot"
/lab_host="TJCI21"
/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI"

BASE COUNT

120 a 167 c 195 g 116 t

ORIGIN

Query Match

Best Local Similarity 3.0%; Score 18; DB 151; Length 598;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 416 cgtcacgactaccgcc 433

|||||

Db 254 CGTCACGACTACCGCC 271

RESULT 28

AU101340
 LOCUS AU101340 618 bp mRNA EST 22-AUG-2000
 DEFINITION AU101340 Rice panicle at ripening stage Oryza sativa cDNA clone
 E1131, mRNA sequence.
 ACCESSION AU101340
 VERSION AU101340.1 GI:9865590
 KEYWORDS EST.
 SOURCE Oryza sativa.
 ORGANISM Oryza sativa.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 618)
 Sasaki, T. and Yamamoto, K.
 Rice cDNA from panicle at ripening stage (2000)
 Unpublished (2000)
 Contact: Takuji Sasaki
 National Institute of Agrobiological Resources
 Rice Genome Research Program
 2-1-2 Kannondai, Tsukuba
 Ibaraki,
 Japan 305
 Tel.: 0298-38-7441
 Fax: 0298-38-7468
 Email: tsasaki@abrr.affrc.go.jp
 PROJECT = 'RGP'.
 E1131_42.

FEATURES
 source
 1..618
 /organism="Oryza sativa"
 /strain="Nipponbare"
 /db_xref="taxon:4530"
 /clone="E1131"
 /clone_lib="Rice panicle at ripening stage"
 /dev_stage="ripening stage"
 /note="Organ: panicle; Rice cDNA from panicle at ripening stage"

BASE COUNT 151 a 166 c 145 g 155 t 1 others
 ORIGIN

Query Match 3.0%; Score 18; DB 107; Length 618;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 394 acgacgaggtctacggcg 411
 ||||||||||||||||
 Db 190 ACGACGAGCTCTACGCG 207

RESULT 29
 BE291467 682 bp mRNA EST 13-JUL-2000
 LOCUS BE291467
 DEFINITION 601085211F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:349371 5',
 mRNA sequence.
 ACCESSION BE291467
 VERSION BE291467.1 GI:9173211
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 682)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1AM8556 row: e column: 04
 High quality sequence stop: 577.
 Location/Qualifiers
 1..682
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:349371"
 /clone_lib="NCI_CGAP_Mam6"
 /sex="female, virgin"
 /tissue_type="infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 215 a 150 c 179 g 138 t
 ORIGIN

Query Match 3.0%; Score 18; DB 165; Length 682;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 543 aacagagagaagaagc 560
 ||||||||||||||||
 Db 337 AACAGAGAGAGAAAGGC 354

RESULT 30
 BG323623 707 bp mRNA EST 27-FEB-2001
 LOCUS BG323623
 DEFINITION 602422094F1 NIH-MGC_14 Homo sapiens cDNA clone IMAGE:4560167 5',
 mRNA sequence.
 ACCESSION BG323623
 VERSION BG323623.1 GI:13130060
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 707)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DTP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1CM1268 row: 1 column: 24
 High quality sequence stop: 656.
 Location/Qualifiers
 1..707
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4560167"
 /clone_lib="NIH-MGC_14"
 /tissue_type="renal cell adenocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: kidney; Vector: pORF7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dt priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of

FEATURES
 source

BASE COUNT 182 a 199 c 163 g 162 t 1 others
 ORIGIN

Query Match 3.0%; Score 18; DB 152; Length 707;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 cccctgcattccccctg 247
 ||||||||||||||||
 Db 194 CCCCTGATCACCCTCG 211

RESULT 31

LOCUS A2631126 711 bp DNA GSS 13-DEC-2000
 DEFINITION IM0485G09F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 ACCESSION A2631126
 VERSION A2631126.1 GI:11753316
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamli, C.,
 Islam, H., Longacre, S., Mahmood, M., Meenen, E., Pedersen, T., Reilly,
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.
 and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0485 row: G column: 09
 Seq primer: CGTTGTAACGACGCGCCAGT
 Class: plasmid ends
 High quality sequence stop: 711.

FEATURES

source
 1. /711
 Location/Qualifiers

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0485G09"
 /clone_1lb="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42ny; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of PMD42 (g1147321149b/AP129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to

BASE COUNT 210 a 180 c 118 g 203 t
 ORIGIN

Query Match 3.0%; Score 18; DB 247; Length 711;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 547 ggaagaagaaggctcc 564
 ||||||||||||||||
 Db 146 GGAGAAGAAAAGCCTCC 163

RESULT 32

LOCUS A2867529/c 712 bp DNA GSS 21-FEB-2001
 DEFINITION 2M0178H23F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 ACCESSION A2867529
 VERSION A2867529.1 GI:13069921
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamli, C.,
 Islam, H., Longacre, S., Mahmood, M., Meenen, E., Pedersen, T., Reilly,
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.
 and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0178 row: H column: 23
 Seq primer: CGTTGTAACGACGCGCCAGT
 Class: plasmid ends
 High quality sequence stop: 712.

FEATURES

source
 1. /712
 Location/Qualifiers

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0178H23"
 /clone_1lb="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42ny; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of PMD42 (g1147321149b/AP129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

BASE COUNT 191 a 213 c 141 g 166 t 1 others

Query Match Best Local Similarity 3.0%; Score 18; DB 251; Length 712;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 297 tgaagccgacgtggggag 314
|||||
Db 525 TGAGCCGACGTCGGGAG 508

RESULT 33

BE892353

LOCUS 722 bp mRNA EST 20-OCT-2000
DEFINITION 601433812P1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3919130 5',
mRNA sequence.

ACCESSION BE892353

VERSION BE892353.1 GI:10352576
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1

(bases 1 to 722)
NIH-MGC http://mgs.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: ggapbs@mail.nih.gov

Tissue Procurement: ATCC/DCTP/DRP
CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov
Plate: LLM9748 row: 9 column: 03

High quality sequence stop: 709.
Location/Qualifiers

1..722
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3919130"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 197 a 198 c 161 g 166 t

ORIGIN

Query Match 3.0%; Score 18; DB 141; Length 722;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 cccctgacacccctg 247
|||||
Db 337 CCCCTGCATCACCCCTG 354

RESULT 34

BF267443 862 bp mRNA EST 09-MAR-2001
LOCUS HV_CEA0017P07f Hordeum vulgare seedling green leaf EST library
DEFINITION

BASE COUNT 193 a 380 c 206 g 80 t 3 others

ORIGIN

Query Match 3.0%; Score 18; DB 146; Length 862;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 ggcgcgcacgacgag 402
|||||
Db 52 GCGCGCGCAGCAGAG 35

RESULT 35

AQ901727/c 898 bp DNA GSS 10-NOV-1999
LOCUS HS_2019_A2_E01_T7C CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate-2019 Col-2 Row-1, DNA sequence.

ACCESSION AQ901727.1 GI:6357917
VERSION AQ901727.1
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 898)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
Contact: Mahairas GG, Wallace JC, Hood L

HVCDNA0004 (Erysiphe infected & control) Hordeum vulgare cDNA clone
HV_CEA0017P07f, mRNA sequence.

ACCESSION BF267443
VERSION BF267443.2 GI:13263387
KEYWORDS EST.

SOURCE barley.
ORGANISM Hordeum vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 862)
Wing,R., Close,T.J., Kleinholz,A., Wise,R., Begum,D., Frisch,D., Yu
Y., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo
T., Saski,C., Schwartzbeck,J., Simmons,J., Choi,D.W., Main,D. and
Wood,T.

Development of a genetically and physically anchored EST resource
for barley genomics
Unpublished (2000)
On Nov 17, 2000 this sequence version replaced gi:11198438.

TITLE Development of a genetically and physically anchored EST resource
JOURNAL Unpublished (2000)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel.: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu

Seq primer: AATTACCTCTACCTAAAGG
High quality sequence stop: 741.
Location/Qualifiers

1..862
/organism="Hordeum vulgare"
/cultivar="Cili6155 (M1a13)"
/db_xref="taxon:4513"
/clone="HV_CEA0017P07f"
/clone_lib="Hordeum vulgare seedling green leaf EST
library HVCDNA0004 (Erysiphe infected & control)"
/tissue_type="seedling green leaf"
/lab_host="TJC121"
/note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 193 a 380 c 206 g 80 t 3 others

ORIGIN

Query Match 3.0%; Score 18; DB 146; Length 862;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 ggcgcgcacgacgag 402
|||||
Db 52 GCGCGCGCAGCAGAG 35

RESULT 35

AQ901727/c 898 bp DNA GSS 10-NOV-1999
LOCUS HS_2019_A2_E01_T7C CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate-2019 Col-2 Row-1, DNA sequence.

ACCESSION AQ901727.1 GI:6357917
VERSION AQ901727.1
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 898)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
Contact: Mahairas GG, Wallace JC, Hood L

BASE COUNT 193 a 380 c 206 g 80 t 3 others

ORIGIN

Query Match 3.0%; Score 18; DB 146; Length 862;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 ggcgcgcacgacgag 402
|||||
Db 52 GCGCGCGCAGCAGAG 35

RESULT 35

AQ901727/c 898 bp DNA GSS 10-NOV-1999
LOCUS HS_2019_A2_E01_T7C CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate-2019 Col-2 Row-1, DNA sequence.

ACCESSION AQ901727.1 GI:6357917
VERSION AQ901727.1
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 898)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
Contact: Mahairas GG, Wallace JC, Hood L

BASE COUNT 193 a 380 c 206 g 80 t 3 others

ORIGIN

Query Match 3.0%; Score 18; DB 146; Length 862;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 ggcgcgcacgacgag 402
|||||
Db 52 GCGCGCGCAGCAGAG 35

RESULT 35

AQ901727/c 898 bp DNA GSS 10-NOV-1999
LOCUS HS_2019_A2_E01_T7C CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate-2019 Col-2 Row-1, DNA sequence.

ACCESSION AQ901727.1 GI:6357917
VERSION AQ901727.1
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 898)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
Contact: Mahairas GG, Wallace JC, Hood L

BASE COUNT 193 a 380 c 206 g 80 t 3 others

ORIGIN

Query Match 3.0%; Score 18; DB 146; Length 862;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 ggcgcgcacgacgag 402
|||||
Db 52 GCGCGCGCAGCAGAG 35

RESULT 35

AQ901727/c 898 bp DNA GSS 10-NOV-1999
LOCUS HS_2019_A2_E01_T7C CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate-2019 Col-2 Row-1, DNA sequence.

High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: <http://www.htsc.washington.edu>
Plate: 2019 row: 1 column: 2
Seq primer: T7
Class: BAC ends
High quality sequence stop: 898.

FEATURES
source
1. .898
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=2019 Col=2 Row=1"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

BASE COUNT 154 a 251 c 168 g 322 t 3 others

ORIGIN

Query Match 3.0%; Score 18; DB 235; Length 898;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 542 caacagagagagagagag 559
|||||

Db 881 CACACGAGACAGACAGAG 864

RESULT 36
CNS01XEF 909 bp DNA GSS 12-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
DEFINITION 202608 of library G from Tetraodon nigroviridis, genomic survey
sequence.
AL171604
ACCESSION AL171604.1 GI:7809661
VERSION GSS: genome survey sequence.
KEYWORDS Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis.
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 909)
Roest-Crollius, H., Jallion, O., Dasilva, C., Fizames, C., Fisher, C.,
Bonneau, L., Billault, A., Quetier, F., Saulin, W., Bernot, A. and
Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Unpublished
2 (bases 1 to 909)
Roest-Crollius, H., Jallion, O., Dasilva, C., Fizames, C., Fisher, C.,
Bonneau, L., Billault, A., Quetier, F., Saulin, W., Bernot, A.,
Saulin, W. and Weissenbach, J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
3 (bases 1 to 909)
Genoscope.
/db_xref="taxon:99883"
/clone="047C14"
/clone_lib="G"
/note="Genoscope sequence ID : C0BG047BB07LP1-end : T7"

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
This sequence is a single read and was generated as part of a large
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.

FEATURES
source
1. .909
Location/Qualifiers

/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="202608"
/clone_lib="G"
/note="Genoscope sequence ID : C0BG202BD04SP1-end :
PUC-ori"

BASE COUNT 192 a 228 c 218 g 258 t 13 others

ORIGIN

Query Match 3.0%; Score 18; DB 220; Length 909;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 acagcagcagccagcaaa 131
|||||

Db 524 ACACGACGCCAGCCACAAA 507

RESULT 37
CNS030F7 919 bp DNA GSS 17-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 047C14 of library G from Tetraodon nigroviridis, genomic survey
sequence.
AL255868
ACCESSION AL255868.1 GI:7976880
VERSION GSS: genome survey sequence.
KEYWORDS Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis.
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 919)
Roest-Crollius, H., Jallion, O., Dasilva, C., Fizames, C., Fisher, C.,
Bonneau, L., Billault, A., Quetier, F., Saulin, W., Bernot, A. and
Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Unpublished
2 (bases 1 to 919)
Roest-Crollius, H., Jallion, O., Dasilva, C., Fizames, C., Fisher, C.,
Bonneau, L., Billault, A., Quetier, F., Saulin, W., Bernot, A.,
Saulin, W. and Weissenbach, J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
3 (bases 1 to 919)
Genoscope.
/db_xref="taxon:99883"
/clone="047C14"
/clone_lib="G"
/note="Genoscope sequence ID : C0BG047BB07LP1-end : T7"

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
This sequence is a single read and was generated as part of a large
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.

FEATURES
source
1. .919
Location/Qualifiers

BASE COUNT 170 a 312 c 280 g 148 t 9 others

ORIGIN

Query Match 3.0%; Score 18; DB 221; Length 919;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 268 gccgcgcctcctgcgcgcg 285
|||||

Db 575 GCCGCGCTCTCTGCGCG 592

RESULT 38
CNS05JPS/C
LOCUS
DEFINITION CNS05JPS 952 bp DNA GSS 26-MAY-2000
Tetraodon nigroviridis genome survey sequence T3 end of clone
028E23 of library A from Tetraodon nigroviridis, genomic survey
sequence.

ACCESSION
VERSION AL340489.1 GI:8234247
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorphae; Acanthopterygii; Percomorphae; Tetraodontiformes;
Tetraodontidae; Tetraodon.

REFERENCE
AUTHORS Roest-Crollius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
Bonneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 952)
AUTHORS Roest-Crollius,H., Jallion,O., Dasilva,C., Bonneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 952)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
Location/Qualifiers
1..952
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="028E23"
/clone_1lb="A"
/note="Genoscope sequence ID : C0AA028AC12A2-end : T3"

BASE COUNT 239 a 231 c 236 g 236 t 10 others

ORIGIN

Query Match 3.0%; Score 18; DB 222; Length 952;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 114 acagcagccagggccaaa 131
|||||
Db 137 ACAGCAGCCAGGCCAAA 120

RESULT 39
BE741337
LOCUS
DEFINITION BE741337 968 bp mRNA EST 15-SEP-2000
601594184F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3947972.5,
mRNA sequence.
ACCESSION BE741337
VERSION BE741337.1 GI:10155329
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 968)

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rt@mail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LRCM810 row: h column: 21
High quality sequence stop: 697.
Location/Qualifiers
1..968
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3947972"
/clone_1lb="NIH_MGC.9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACAGC(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the Laboratory of Gerald M. Rubin (University of
California, Berkeley) using 2AP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 253 a 275 c 237 g 203 t

ORIGIN

Query Match 3.0%; Score 18; DB 139; Length 968;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 230 cccctgacaccacctg 247
|||||
Db 277 CCCCTGCATCACCCTG 294

RESULT 40
CNS035RQ
LOCUS
DEFINITION CNS035RQ 966 bp DNA GSS 15-MAY-2000
Tetraodon nigroviridis genome survey sequence T7 end of clone
214B05 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL229103.1 GI:7888098
VERSION AL229103
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorphae; Acanthopterygii; Percomorphae; Tetraodontiformes;
Tetraodontidae; Tetraodon.

REFERENCE 1 (bases 1 to 966)
AUTHORS Roest-Crollius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
Bonneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 966)
AUTHORS Roest-Crollius,H., Jallion,O., Dasilva,C., Bonneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 966)
AUTHORS Genoscope.

TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES
Source
1. 986
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone_1lb="G"
/note="Genoscope sequence ID : COAG214CA03LP1-end : T7"
BASE COUNT 180 a 329 c 318 g 154 t 5 others
ORIGIN

Query Match 3.0%; Score 18; DB 220; Length 986;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 268 gccgcgcctcctgcgcgcg 285
|||||
Db 572 GCCGCGCCTCCTCGCGG 589

RESULT 41
CNS0305J 1009 bp DNA GSS 17-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 046L1 of library G from Tetraodon nigroviridis, genomic survey sequence.

ACCESSION AL255520.1 GI:7976532
VERSION AL255520.1
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
1 (bases 1 to 1009)
Roest-Crollius, H., Jallion, O., Dasilva, C., Fizames, C., Fisher, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.
Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
Unpublished
2 (bases 1 to 1009)
Roest-Crollius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brotlier, P., Quetier, F., Saurin, W. and Weissenbach, J.
Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
Unpublished
3 (bases 1 to 1009)
Genoscope.
Direct Submission
Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES
Source
1. 1009
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone_1lb="G"
/note="Genoscope sequence ID : COBG046CF06LP1-end : T7"
BASE COUNT 211 a 263 c 238 g 294 t 3 others
ORIGIN

Query Match 3.0%; Score 18; DB 221; Length 1009;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 acagcagccagcgccaana 131
|||||
Db 736 ACAGCAGCCAGCGCCAAA 719

RESULT 42
CNS05EAA 1056 bp DNA GSS 26-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence T3 end of clone
DEFINITION 028A19 of library A from Tetraodon nigroviridis, genomic survey sequence.

ACCESSION AL333451 GI:8227209
VERSION AL333451
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
1 (bases 1 to 1056)
Roest-Crollius, H., Jallion, O., Dasilva, C., Fizames, C., Fisher, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.
Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
Unpublished
2 (bases 1 to 1056)
Roest-Crollius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brotlier, P., Quetier, F., Saurin, W. and Weissenbach, J.
Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
Unpublished
3 (bases 1 to 1056)
Genoscope.
Direct Submission
Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES
Source
1. 1056
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone_1lb="A"
/note="Genoscope sequence ID : COA028A19A1-end : T3"
BASE COUNT 223 a 246 c 262 g 278 t 47 others
ORIGIN

Query Match 3.0%; Score 18; DB 222; Length 1056;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 acagcagccagcgccaana 131
|||||
Db 148 ACAGCAGCCAGCGCCAAA 131

RESULT 43
AL543701 1078 bp mRNA EST 16-FEB-2001
LOCUS AL543701 L7L NFE006_P12 Homo sapiens cDNA clone CS0D1006YNI0 5
DEFINITION prime, mRNA sequence.
ACCESSION AL543701

| | | |
|-----------------------|---|------------------------------------|
| VERSION | AL543701.1 | GI:12876180 |
| KEYWORDS | ESF. | |
| SOURCE | human. | |
| ORGANISM | Homo sapiens | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| AUTHORS | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | |
| TITLE | 1 (bases 1 to 1078) | |
| JOURNAL | Li, W.B., Gruber, C., Jesse, J. and Polayes, D. | |
| COMMENT | Full-length cDNA libraries and normalization Unpublished (2001) Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr. | |
| FEATURES | Location/Qualifiers | |
| source | 1..1078 | |
| | /organism="Homo sapiens" | |
| | /db_xref="taxon:9606" | |
| | /clone="CS0D1006YN10" | |
| | /clone_lib="LTL_NFL006_PL2" | |
| | /tissue_type="placenta" | |
| | /note="Vector: pCMVSPORT 6; Site: 1: NotI; 1st strand cDNA was primed with a NotI-Oligo(gt) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com" | |
| BASE COUNT | 246 a 287 c 329 g 201 t | 15 others |
| ORIGIN | | |
| Query Match | 3.0% | Score 18; DB 106; Length 1078; |
| Best Local Similarity | 100.0% | Pred. No. 78; |
| Matches | 18; Conservative | 0; Mismatches 0; Indels 0; Gaps 0; |
| QY | 230 cccctgcacacccccctg 247 | |
| DB | 683 CCCGCGATCACCCCTG 700 | |
| RESULT 44 | | |
| CNS054X1/C | | |
| LOCUS | CNS054X1 1081 bp DNA GSS 26-Jul-2000 | |
| DEFINITION | Tetradon nigroviridis genome survey sequence T7 end of clone 046L22 of library A from Tetradon nigroviridis, genomic survey sequence. | |
| ACCESSION | AL321310 | |
| VERSION | AL321310.1 | GI:9554194 |
| KEYWORDS | GSS: genome survey sequence. | |
| SOURCE | Tetradon nigroviridis. | |
| ORGANISM | Tetradon nigroviridis Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetradon. | |
| REFERENCE | 1 (bases 1 to 1081) | |
| AUTHORS | Roest Crolius, H., Jallou, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brothier, P., Quetier, F., Saurin, W. and Weissbach, J. | |
| TITLE | Estimate of human gene number provided by genome-wide analysis using Tetradon nigroviridis DNA sequence | |
| JOURNAL | Nat. Genet. 25 (2), 235-238 (2000) | |
| REFERENCE | 20296633 | |
| MEDLINE | 2 (bases 1 to 1081) | |
| AUTHORS | Crolius, H.R., Jallou, O., Dasilva, C., Ozouf-Costaz, C., Fizames, C., Fischer, C., Bouneau, L., Ballault, A., Quetier, F., Saurin, W., Bernot, A. and Weissbach, J. | |
| TITLE | Characterization and repeat analysis of the compact genome of the freshwater pufferfish tetradon nigroviridis | |

| FEATURES | source |
|-----------------------|---|
| JOURNAL MEDLINE | Genome Res. 10 (7), 939-949 (2000) |
| REFERENCE | 20359837 |
| AUTHORS | 3 (bases 1 to 1081) |
| TITLE | Genoscope. |
| JOURNAL | Direct Submission |
| COMMENT | Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon . |
| FEATURES | Location/Qualifiers |
| source | 1..1081 |
| BASE COUNT | 269 a 256 c 267 g 282 t 7 others |
| ORIGIN | |
| Query Match | 3.0%; Score 18; DB 222; Length 1081; |
| Best Local Similarity | 100.0%; Pred. No. 78; |
| Matches | 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| OR | 114 acagcagccagcgccaaa 131 Db 123 ACAGCAGCCAGCCAAA 106 |
| RESULT 45 | |
| AK006925 | |
| LOCUS | AK006925 1101 bp mRNA HTC 08-FEB-2001 |
| DEFINITION | Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:1700069015, full insert sequence. |
| ACCESSION | AK006925 |
| VERSION | AK006925.1 GI:12840231 |
| KEYWORDS | CAP trapper. |
| SOURCE | Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone_11b:RIKEN full-length enriched mouse cDNA library clone:1700069015. |
| ORGANISM | Mus musculus |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |
| AUTHORS | Carninci, P. and Hayashizaki, Y. |
| TITLE | High-efficiency full-length cDNA cloning |
| JOURNAL | Methods Enzymol. 303, 19-44 (1999) |
| REFERENCE | 2 (sites) |
| AUTHORS | Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. |
| TITLE | Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes |
| JOURNAL | Genome Res. 10 (10), 1617-1630 (2000) |
| MEDLINE | 20499374 |
| REFERENCE | 3 (sites) |
| AUTHORS | Kono, H., Akiyama, J., Nishik, K., Kitsu, T., Tashiro, H., Itoh, M., Kikuchi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiyagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsuhira, M., Onoda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Watanabe, S., Okazaki, Y., Muramatsu, M., Inoue, Y., and Hayashizaki, Y. |
| TITLE | RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer |
| JOURNAL | Genome Res. 10 (11), 1757-1771 (2000) |
| MEDLINE | 20530913 |
| REFERENCE | 4 (sites) |
| AUTHORS | The RIKEN Genome Exploration Research Group Phase II Team and FANTOM Consortium. |
| TITLE | Functional annotation of a full-length mouse cDNA collection - Nature 409, 685-690 (2001) |
| JOURNAL | 5 (bases 1 to 1101) |

Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10289 row: j column: 10
High quality sequence stop: 28.

FEATURES

Source

1. 1481
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4471017"
/clone_1ib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Liver; Vector: PCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library."

BASE COUNT 460 a 355 c 326 g 340 t

ORIGIN

Query Match 3.0%; Score 18; DB 175; Length 1481;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 539 aaccacagagagaagaa 556
|||||
DB 1371 AACCAACAGACAGAGAA 1388

RESULT 48

A1701300/c

LOCUS A1701300 164 bp mRNA EST 17-DEC-1999
DEFINITION w559h07.x1 NCI CGAP Panl Homo sapiens cDNA clone IMAGE:2322973 3'
similar to gb:M1606 PHOSPHORYLASE B KINASE GAMMA CATALYTIC CHAIN,
TESTIS (HOMAN);, mRNA sequence.

ACCESSION A1701300 GI:4989200
VERSION A1701300.1
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).

Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Life Technologies catalog #: 11548-013

DNA Sequencing by: Washington University Genome Sequencing Center
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1026 Std Error: 0.00
Seq primer: -40UP from GIBCO
High quality sequence stop: 1.

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

FEATURES

Source

1. 164
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2322973"
/clone_1ib="NCI CGAP Panl"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: PCMV-SPORT6; Site: 1: SalI;
Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"
BASE COUNT 34 a 40 c 53 g 36 t 1 others

ORIGIN

Query Match 2.9%; Score 17; DB 23; Length 164;
Best Local Similarity 100.0%; Pred. No. 2,7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 222 tgcgtacccctgcacat 238
|||||
DB 43 TCGTCATCCCTGCAT 27

RESULT 49

AA002343

LOCUS AA002343 165 bp mRNA EST 19-JUL-1996
DEFINITION mg42a03.r1 Soares mouse embryo NBME13.5 14.5 Mus musculus cDNA
clone IMAGE:426412 5', mRNA sequence.

ACCESSION AA002343 GI:1445745
VERSION AA002343.1
KEYWORDS EST.

SOURCE house mouse.
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 165)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Gelsel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

FEATURES

The WashU-HMT Mouse EST Project
Unpublished (1996)
Contact: Maria M/Mouse EST Project
WashU-HMT Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (infoimage.llnl.gov) for further information.
MGI:260964
Seq primer: mob.REGA+ET
High quality sequence stop: 156.
Location/Qualifiers

Source

1. 165
Location/Qualifiers

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:426412"
/clone_1ib="Soares mouse embryo NBME13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"

/note="Vector: pRTT3D-Pac (Pharmacia) with a modified
polylinker; Site: 1: Not I; Site: 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer 15'
TGTATCAATCTGAGAGCGGCGCGCGGAAATTTTCTTTTCTTTTCTTTT
T 3'1, on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos (total RNA provided by Minoru Ko, Wayne
State Univ., from 2 j; double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pRTT3 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M. Fatima Bonaldo."

BASE COUNT 52 a 38 c 47 g 28 t

ORIGIN

Query Match 2.9%; Score 17; DB 1; Length 165;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 aacagcagccagcccaa 129
 |||||||||||||||
 Db 50 AACAGCAGCCAGGCCAA 66

QY 113 aacagcagccagcccaa 129
 |||||||||||||||
 Db 115 AACAGCAGCCAGGCCAA 131

Search completed: September 21, 2001, 03:16:32
 Job time: 9108 sec

RESULT 50
 AA042040
 LOCUS AA042040 206 bp mRNA EST 03-SEP-1996
 DEFINITION mJ06a07.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
 clone IMAGE:475284 5', mRNA sequence.
 ACCESSION AA042040
 VERSION AA042040.1 GI:1520230
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
 1 (bases 1 to 206)
 Mairra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 The WashU-HHMI Mouse EST Project
 Unpublished (1996)
 Contact: Mairra M/Mouse EST Project
 WashU-HHMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:286028
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 151.
 Location/Qualifiers
 1..206
 /organism="Mus musculus"
 /strain="C57Bl/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:475284"
 /clone_1lb="Soares mouse embryo NbME13.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - 01190(dt) primer [5'
 TGTTCACATCGAAGCGAGCGCGCGCAAAATTTTTTTTTTTTTTTT
 T 3']; on equal amounts of RNA from 2 13.5dpc and 2
 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
 State Univ., from 2]; double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT73 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M.Fatima Bonaldo."

BASE COUNT 52 a 47 c 68 g 39 t
 ORIGIN

Query Match 2.9%; Score 17; DB 1; Length 206;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 13:57:09 : Search time 84.41 Seconds
(without alignments)
1332.198 Million cell updates/sec

Title: US-09-138-735-1_COPY_1232_1825

Perfect score: 594

Sequence: 1 caggtacacgacgtacgactt.....aaacgagctcgtgcgata 594

Scoring table:

OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 324599 seqs, 94655562 residues

Word size : 0

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

Issued_Patents_MA:*
1: /cgnl_7/ptodata/1/ina/5A.COMB.seq:*
2: /cgnl_7/ptodata/1/ina/5B.COMB.seq:*
3: /cgnl_7/ptodata/1/ina/6A.COMB.seq:*
4: /cgnl_7/ptodata/1/ina/6B.COMB.seq:*
5: /cgnl_7/ptodata/1/ina/PCrUS.COMB.seq:*
6: /cgnl_7/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1 | 594 | 100.0 | 3402 | 1 US-08-480-917-1 | Sequence 1, Appli |
| 2 | 22 | 3.7 | 22 | 1 US-08-480-917-9 | Sequence 9, Appli |
| 3 | 19 | 3.2 | 11873 | 2 US-08-970-269A-32 | Sequence 32, Appli |
| 4 | 19 | 3.2 | 11878 | 2 US-08-970-269A-31 | Sequence 31, Appli |
| 5 | 19 | 3.2 | 11883 | 2 US-08-970-269A-28 | Sequence 28, Appli |
| 6 | 18 | 3.0 | 18 | 1 US-08-480-917-5 | Sequence 5, Appli |
| 7 | 18 | 3.0 | 18 | 1 US-08-480-917-8 | Sequence 8, Appli |
| 8 | 18 | 3.0 | 2520 | 2 US-08-454-557C-50 | Sequence 50, Appli |
| 9 | 18 | 3.0 | 2520 | 2 US-08-340-426D-50 | Sequence 50, Appli |
| 10 | 18 | 3.0 | 2520 | 2 US-08-450-673C-50 | Sequence 50, Appli |
| 11 | 18 | 3.0 | 2520 | 5 PCT-US95-17111A-50 | Sequence 50, Appli |
| 12 | 18 | 3.0 | 68750 | 3 US-09-335-409-1 | Sequence 1, Appli |
| 13 | 17 | 2.9 | 673 | 2 US-08-692-787-4 | Sequence 4, Appli |
| 14 | 17 | 2.9 | 673 | 4 US-09-097-199-4 | Sequence 4, Appli |
| 15 | 17 | 2.9 | 3337 | 4 US-08-072-610-1 | Sequence 1, Appli |
| 16 | 17 | 2.9 | 3337 | 4 US-08-719-822B-1 | Sequence 1, Appli |
| 17 | 17 | 2.9 | 3337 | 4 US-09-092-458-1 | Sequence 1, Appli |
| 18 | 16 | 2.7 | 41 | 1 US-08-142-887-1 | Sequence 1, Appli |
| 19 | 16 | 2.7 | 1185 | 2 US-08-628-039-9 | Sequence 9, Appli |
| 20 | 16 | 2.7 | 1185 | 2 US-08-912-205-9 | Sequence 9, Appli |
| 21 | 16 | 2.7 | 1185 | 4 US-09-440-400-9 | Sequence 9, Appli |
| 22 | 16 | 2.7 | 1314 | 2 US-09-031-059-2 | Sequence 2, Appli |
| 23 | 16 | 2.7 | 1411 | 1 US-07-952-817-10 | Sequence 10, Appli |
| 24 | 16 | 2.7 | 1411 | 6 5210025-3 | Patent No. 5210025 |
| 25 | 16 | 2.7 | 1856 | 2 US-08-360-606B-29 | Sequence 29, Appli |
| 26 | 16 | 2.7 | 1864 | 2 US-08-673-388-9 | Sequence 9, Appli |
| 27 | 16 | 2.7 | 1864 | 2 US-08-614-877-9 | Sequence 9, Appli |

| | | | | | |
|-----|----|-----|--------|----------------------|--------------------|
| 28 | 16 | 2.7 | 2214 | 3 US-08-864-038A-1 | Sequence 1, Appli |
| 29 | 16 | 2.7 | 2310 | 3 US-09-036-987A-25 | Sequence 25, Appli |
| 30 | 16 | 2.7 | 3331 | 3 US-08-864-038A-2 | Sequence 2, Appli |
| 31 | 16 | 2.7 | 3331 | 3 US-08-864-038A-4 | Sequence 4, Appli |
| 32 | 16 | 2.7 | 5117 | 5 PCT-US95-05512-1 | Sequence 1, Appli |
| 33 | 16 | 2.7 | 5117 | 5 PCT-US95-13749-3 | Sequence 1, Appli |
| 34 | 16 | 2.7 | 6545 | 5 US-08-056-200-03 | Sequence 3, Appli |
| 35 | 16 | 2.7 | 9551 | 2 US-08-800-644-83 | Sequence 93, Appli |
| 36 | 16 | 2.7 | 9551 | 2 US-08-633-779-2 | Sequence 2, Appli |
| 37 | 15 | 2.5 | 60 | 1 US-08-238-883-14 | Sequence 14, Appli |
| 38 | 15 | 2.5 | 81 | 1 US-08-443-407-14 | Sequence 14, Appli |
| 39 | 15 | 2.5 | 81 | 5 PCT-US95-05600-158 | Sequence 158, App |
| 40 | 15 | 2.5 | 325 | 3 US-08-994-946A-1 | Sequence 1, Appli |
| 41 | 15 | 2.5 | 837 | 4 US-08-123-492A-3 | Sequence 1, Appli |
| 42 | 15 | 2.5 | 837 | 4 US-08-633-779-1 | Sequence 1, Appli |
| 43 | 15 | 2.5 | 837 | 4 US-09-154-874-1 | Sequence 1, Appli |
| 44 | 15 | 2.5 | 846 | 3 US-09-082-092-8 | Sequence 8, Appli |
| 45 | 15 | 2.5 | 1281 | 4 US-08-176-427B-7 | Sequence 7, Appli |
| 46 | 15 | 2.5 | 1313 | 2 US-08-356-060A-4 | Sequence 4, Appli |
| 47 | 15 | 2.5 | 1313 | 4 US-08-460-900C-4 | Sequence 4, Appli |
| 48 | 15 | 2.5 | 1338 | 4 US-08-855-910-5 | Sequence 5, Appli |
| 49 | 15 | 2.5 | 1423 | 3 US-08-651-136C-9 | Sequence 9, Appli |
| 50 | 15 | 2.5 | 1608 | 2 US-07-959-941-1 | Sequence 8, Appli |
| 51 | 15 | 2.5 | 1632 | 1 US-08-259-924-1 | Sequence 1, Appli |
| 52 | 15 | 2.5 | 1632 | 1 US-08-757-230A-1 | Sequence 1, Appli |
| 53 | 15 | 2.5 | 1715 | 5 PCT-US95-02315-1 | Sequence 1, Appli |
| 54 | 15 | 2.5 | 1715 | 5 US-09-348-443-1 | Sequence 1, Appli |
| 55 | 15 | 2.5 | 1763 | 3 US-08-934-494-1 | Sequence 1, Appli |
| 56 | 15 | 2.5 | 1869 | 3 US-09-143-068-1 | Sequence 1, Appli |
| 57 | 15 | 2.5 | 1869 | 4 US-09-350-268-1 | Sequence 1, Appli |
| 58 | 15 | 2.5 | 1869 | 4 US-08-642-406A-21 | Sequence 21, Appli |
| 59 | 15 | 2.5 | 1875 | 2 US-09-082-092-5 | Sequence 5, Appli |
| 60 | 15 | 2.5 | 1876 | 4 US-09-227-717-3 | Sequence 3, Appli |
| 61 | 15 | 2.5 | 2044 | 4 US-08-787-739-58 | Sequence 58, Appli |
| 62 | 15 | 2.5 | 2501 | 6 5457037-4 | Patent No. 5457037 |
| 63 | 15 | 2.5 | 2625 | 6 US-08-365-189-7 | Sequence 7, Appli |
| 64 | 15 | 2.5 | 2646 | 1 US-08-365-189-2 | Sequence 2, Appli |
| 65 | 15 | 2.5 | 2982 | 1 US-08-365-189-1 | Sequence 1, Appli |
| 66 | 15 | 2.5 | 2982 | 1 US-08-826-246-3 | Sequence 3, Appli |
| 67 | 15 | 2.5 | 3103 | 3 US-08-944-495-3 | Sequence 3, Appli |
| 68 | 15 | 2.5 | 3103 | 3 US-09-126-640-2 | Sequence 2, Appli |
| 69 | 15 | 2.5 | 3103 | 4 US-08-925-588-3 | Sequence 3, Appli |
| 70 | 15 | 2.5 | 3103 | 4 US-09-487-444-3 | Sequence 3, Appli |
| 71 | 15 | 2.5 | 3111 | 3 5457037-1 | Patent No. 5457037 |
| 72 | 15 | 2.5 | 3336 | 6 US-08-642-406A-20 | Sequence 20, Appli |
| 73 | 15 | 2.5 | 3517 | 3 US-08-434-000A-1 | Sequence 1, Appli |
| 74 | 15 | 2.5 | 3517 | 3 US-08-787-739-50 | Sequence 90, Appli |
| 75 | 15 | 2.5 | 3532 | 3 US-09-060-618-7 | Sequence 7, Appli |
| 76 | 15 | 2.5 | 3981 | 4 US-09-060-618-7 | Sequence 7, Appli |
| 77 | 15 | 2.5 | 3981 | 4 US-09-060-618-7 | Sequence 7, Appli |
| 78 | 15 | 2.5 | 4296 | 4 US-08-472-358-1 | Sequence 1, Appli |
| 79 | 15 | 2.5 | 4983 | 5 PCT-US92-05786A-1 | Sequence 1, Appli |
| 80 | 15 | 2.5 | 4983 | 5 US-08-485-862B-5 | Sequence 5, Appli |
| 81 | 15 | 2.5 | 4983 | 5 US-08-485-862B-5 | Sequence 5, Appli |
| 82 | 15 | 2.5 | 4984 | 2 US-08-487-077A-5 | Sequence 5, Appli |
| 83 | 15 | 2.5 | 4984 | 2 US-08-485-863A-5 | Sequence 5, Appli |
| 84 | 15 | 2.5 | 10898 | 2 US-08-477-504A-5 | Sequence 5, Appli |
| 85 | 15 | 2.5 | 10898 | 2 US-08-485-862B-5 | Sequence 5, Appli |
| 86 | 15 | 2.5 | 10898 | 2 US-08-485-862B-5 | Sequence 5, Appli |
| 87 | 15 | 2.5 | 10898 | 3 US-08-487-077A-5 | Sequence 5, Appli |
| 88 | 15 | 2.5 | 10898 | 3 US-08-485-863A-5 | Sequence 5, Appli |
| 89 | 15 | 2.5 | 10898 | 3 US-08-485-863A-5 | Sequence 5, Appli |
| 90 | 15 | 2.5 | 10898 | 3 US-08-485-863A-5 | Sequence 5, Appli |
| 91 | 15 | 2.5 | 24417 | 2 US-08-846-762-1 | Sequence 1, Appli |
| 92 | 15 | 2.5 | 50341 | 1 US-08-247-901C-1 | Sequence 1, Appli |
| 93 | 15 | 2.5 | 50341 | 2 US-09-075-904-1 | Sequence 1, Appli |
| 94 | 15 | 2.5 | 50937 | 4 US-09-428-517-1 | Sequence 1, Appli |
| 95 | 15 | 2.5 | 52297 | 4 US-09-428-517-1 | Sequence 1, Appli |
| 96 | 15 | 2.5 | 246240 | 2 US-08-724-394A-20 | Sequence 20, Appli |
| 97 | 15 | 2.5 | 246240 | 2 US-08-724-394A-21 | Sequence 21, Appli |
| 98 | 15 | 2.5 | 246240 | 2 US-08-724-394A-22 | Sequence 22, Appli |
| 99 | 14 | 2.4 | 30 | 4 US-09-481-288-4 | Sequence 4, Appli |
| 100 | 14 | 2.4 | 48 | 4 US-08-916-576B-31 | Sequence 31, Appli |

| | | | | | | | | | | | | | |
|-------|----|-----|------|---|--------------------|--------------------|-----|----|-----|------|---|--------------------|--------------------|
| C 101 | 14 | 2.4 | 213 | 4 | US-08-905-223-46 | Sequence 46, App1 | 174 | 14 | 2.4 | 1392 | 3 | US-08-850-227-3 | Sequence 3, App11 |
| C 102 | 14 | 2.4 | 282 | 6 | 5218099-8 | Patent No. 5218099 | 175 | 14 | 2.4 | 1392 | 3 | US-08-850-227-6 | Sequence 6, App11 |
| C 103 | 14 | 2.4 | 297 | 2 | US-08-637-759B-19 | Sequence 19, App1 | 176 | 14 | 2.4 | 1392 | 3 | US-09-054-985A-3 | Sequence 3, App11 |
| C 104 | 14 | 2.4 | 297 | 3 | US-08-871-355A-21 | Sequence 19, App1 | 176 | 14 | 2.4 | 1392 | 3 | US-09-054-985A-6 | Sequence 6, App11 |
| C 105 | 14 | 2.4 | 298 | 2 | US-08-637-759B-19 | Sequence 21, App1 | 177 | 14 | 2.4 | 1392 | 3 | US-08-812-411-1 | Sequence 1, App11 |
| C 106 | 14 | 2.4 | 298 | 3 | US-08-871-355A-21 | Sequence 21, App1 | 178 | 14 | 2.4 | 1401 | 2 | US-08-812-411-1 | Sequence 1, App11 |
| C 107 | 14 | 2.4 | 326 | 3 | US-08-916-576B-16 | Sequence 16, App1 | 179 | 14 | 2.4 | 1401 | 4 | US-09-180-271-4 | Sequence 4, App11 |
| C 108 | 14 | 2.4 | 326 | 4 | US-08-916-576B-16 | Sequence 16, App1 | 180 | 14 | 2.4 | 1417 | 1 | US-08-713-828-2 | Sequence 2, App11 |
| C 109 | 14 | 2.4 | 399 | 3 | US-08-894-699-48 | Sequence 18, App1 | 181 | 14 | 2.4 | 1417 | 2 | US-08-919-627-2 | Sequence 2, App11 |
| C 110 | 14 | 2.4 | 399 | 3 | US-08-894-699-48 | Sequence 58, App1 | 182 | 14 | 2.4 | 1417 | 2 | US-09-096-245-1 | Sequence 2, App11 |
| C 111 | 14 | 2.4 | 399 | 3 | US-08-894-699-48 | Sequence 59, App1 | 183 | 14 | 2.4 | 1417 | 2 | US-08-875-972-1 | Sequence 1, App11 |
| C 112 | 14 | 2.4 | 421 | 6 | 5196523-22 | Sequence 61, App1 | 184 | 14 | 2.4 | 1423 | 4 | US-08-916-576B-3 | Sequence 3, App11 |
| C 113 | 14 | 2.4 | 435 | 6 | 5169835-7 | Patent No. 5169835 | 185 | 14 | 2.4 | 1437 | 4 | US-08-252-966B-14 | Sequence 14, App1 |
| C 114 | 14 | 2.4 | 483 | 1 | US-08-592-126-118 | Sequence 118, App | 186 | 14 | 2.4 | 1498 | 3 | US-09-173-581-14 | Sequence 14, App1 |
| C 115 | 14 | 2.4 | 501 | 2 | US-08-611-757-37 | Sequence 37, App1 | 187 | 14 | 2.4 | 1518 | 2 | US-08-997-080-88 | Sequence 88, App1 |
| C 116 | 14 | 2.4 | 505 | 5 | PCT-US95-05980-37 | Sequence 37, App1 | 188 | 14 | 2.4 | 1518 | 2 | US-08-997-362-88 | Sequence 88, App1 |
| C 117 | 14 | 2.4 | 505 | 4 | US-08-916-576B-18 | Sequence 18, App1 | 189 | 14 | 2.4 | 1518 | 3 | US-08-873-970-88 | Sequence 88, App1 |
| C 118 | 14 | 2.4 | 512 | 2 | US-08-481-658B-45 | Sequence 45, App1 | 190 | 14 | 2.4 | 1518 | 4 | US-09-095-855-88 | Sequence 88, App1 |
| C 119 | 14 | 2.4 | 512 | 2 | US-08-477-504A-45 | Sequence 45, App1 | 191 | 14 | 2.4 | 1536 | 4 | US-09-318-443-5 | Sequence 5, App11 |
| C 120 | 14 | 2.4 | 512 | 2 | US-08-477-504A-45 | Sequence 45, App1 | 192 | 14 | 2.4 | 1591 | 2 | US-08-194-987B-3 | Sequence 3, App11 |
| C 121 | 14 | 2.4 | 512 | 2 | US-08-486-756A-45 | Sequence 45, App1 | 193 | 14 | 2.4 | 1601 | 3 | US-08-817-188-3 | Sequence 3, App11 |
| C 122 | 14 | 2.4 | 512 | 2 | US-08-485-862B-45 | Sequence 45, App1 | 194 | 14 | 2.4 | 1633 | 1 | US-07-866-979-5 | Sequence 5, App11 |
| C 123 | 14 | 2.4 | 512 | 3 | US-08-787-739-45 | Sequence 45, App1 | 195 | 14 | 2.4 | 1633 | 2 | US-08-466-906B-5 | Sequence 5, App11 |
| C 124 | 14 | 2.4 | 512 | 3 | US-08-485-863A-45 | Sequence 45, App1 | 196 | 14 | 2.4 | 1633 | 3 | US-08-706-281A-5 | Sequence 5, App11 |
| C 125 | 14 | 2.4 | 512 | 3 | US-08-485-863A-45 | Sequence 45, App1 | 197 | 14 | 2.4 | 1633 | 4 | US-09-201-746-5 | Sequence 5, App11 |
| C 126 | 14 | 2.4 | 522 | 4 | US-08-485-863A-45 | Sequence 45, App1 | 198 | 14 | 2.4 | 1640 | 3 | US-09-330-095-2 | Sequence 3, App11 |
| C 127 | 14 | 2.4 | 522 | 4 | US-08-906-136A-3 | Sequence 3, App11 | 199 | 14 | 2.4 | 1659 | 2 | US-08-522-421-1 | Sequence 1, App11 |
| C 128 | 14 | 2.4 | 525 | 3 | US-08-906-136A-1 | Sequence 3, App11 | 200 | 14 | 2.4 | 1660 | 2 | US-08-759-581B-3 | Sequence 3, App11 |
| C 129 | 14 | 2.4 | 525 | 4 | US-09-363-427-1 | Sequence 1, App11 | 201 | 14 | 2.4 | 1682 | 4 | US-09-318-443-7 | Sequence 7, App11 |
| C 130 | 14 | 2.4 | 532 | 1 | US-07-872-878A-10 | Sequence 10, App1 | 202 | 14 | 2.4 | 1689 | 4 | US-09-311-924-1 | Sequence 1, App11 |
| C 131 | 14 | 2.4 | 536 | 2 | US-08-705-875A-7 | Sequence 7, App11 | 203 | 14 | 2.4 | 1695 | 3 | US-09-008-481A-3 | Sequence 3, App11 |
| C 132 | 14 | 2.4 | 536 | 2 | US-08-557-128-9 | Sequence 7, App11 | 204 | 14 | 2.4 | 1695 | 3 | US-09-195-666A-17 | Sequence 17, App1 |
| C 133 | 14 | 2.4 | 599 | 2 | US-08-705-875A-2 | Sequence 2, App11 | 205 | 14 | 2.4 | 1733 | 1 | US-08-470-202-58 | Sequence 58, App1 |
| C 134 | 14 | 2.4 | 616 | 4 | US-08-998-416-1055 | Sequence 1055, App | 206 | 14 | 2.4 | 1733 | 2 | US-08-471-770-58 | Sequence 58, App1 |
| C 135 | 14 | 2.4 | 728 | 4 | US-09-276-531-128 | Sequence 128, App | 207 | 14 | 2.4 | 1743 | 3 | US-08-468-059-58 | Sequence 58, App1 |
| C 136 | 14 | 2.4 | 821 | 4 | US-08-990-823-62 | Sequence 62, App1 | 208 | 14 | 2.4 | 1820 | 2 | US-08-482-677-9 | Sequence 2, App11 |
| C 137 | 14 | 2.4 | 939 | 1 | US-08-420-235B-10 | Sequence 10, App1 | 209 | 14 | 2.4 | 1820 | 2 | US-08-828-922-2 | Sequence 2, App11 |
| C 138 | 14 | 2.4 | 939 | 1 | US-08-793-624-10 | Sequence 10, App1 | 210 | 14 | 2.4 | 1821 | 1 | US-08-201-116-10 | Sequence 10, App1 |
| C 139 | 14 | 2.4 | 939 | 5 | PCT-US95-10194-10 | Sequence 10, App1 | 211 | 14 | 2.4 | 1832 | 2 | US-08-238-821B-10 | Sequence 10, App1 |
| C 140 | 14 | 2.4 | 943 | 3 | US-08-705-875A-1 | Sequence 1, App11 | 212 | 14 | 2.4 | 1832 | 5 | PCT-US95-05744-10 | Sequence 10, App1 |
| C 141 | 14 | 2.4 | 943 | 3 | US-08-705-875A-3 | Sequence 3, App11 | 213 | 14 | 2.4 | 1832 | 5 | US-08-201-118-4 | Sequence 4, App11 |
| C 142 | 14 | 2.4 | 951 | 1 | US-08-671-525B-1 | Sequence 1, App11 | 214 | 14 | 2.4 | 1854 | 2 | US-08-338-821B-4 | Sequence 4, App11 |
| C 143 | 14 | 2.4 | 951 | 1 | US-08-672-109B-1 | Sequence 1, App11 | 215 | 14 | 2.4 | 1854 | 5 | PCT-US95-05744-4 | Sequence 4, App11 |
| C 144 | 14 | 2.4 | 951 | 1 | US-08-842-045-1 | Sequence 1, App11 | 216 | 14 | 2.4 | 1859 | 1 | US-08-234-939-4 | Sequence 4, App11 |
| C 145 | 14 | 2.4 | 951 | 2 | US-08-842-238-1 | Sequence 1, App11 | 217 | 14 | 2.4 | 1859 | 1 | US-08-558-865-4 | Sequence 4, App11 |
| C 146 | 14 | 2.4 | 951 | 3 | US-08-629-335B-1 | Sequence 1, App11 | 218 | 14 | 2.4 | 1863 | 2 | US-08-455-073A-3 | Sequence 3, App11 |
| C 147 | 14 | 2.4 | 968 | 3 | US-08-705-875A-9 | Sequence 9, App11 | 219 | 14 | 2.4 | 1928 | 3 | US-09-008-481A-9 | Sequence 9, App11 |
| C 148 | 14 | 2.4 | 971 | 1 | US-08-253-155A-11 | Sequence 11, App1 | 220 | 14 | 2.4 | 1928 | 3 | US-09-008-481A-9 | Sequence 15, App1 |
| C 149 | 14 | 2.4 | 1083 | 4 | US-09-276-531-30 | Sequence 30, App1 | 221 | 14 | 2.4 | 1933 | 1 | US-09-195-666A-15 | Sequence 3, App11 |
| C 150 | 14 | 2.4 | 1089 | 3 | US-09-195-666A-10 | Sequence 10, App1 | 222 | 14 | 2.4 | 1933 | 1 | US-08-234-939-3 | Sequence 3, App11 |
| C 151 | 14 | 2.4 | 1097 | 2 | US-08-907-492A-1 | Sequence 10, App1 | 223 | 14 | 2.4 | 1938 | 6 | US-08-558-865-3 | Sequence 3, App11 |
| C 152 | 14 | 2.4 | 1110 | 1 | US-08-257-341-1 | Sequence 1, App11 | 224 | 14 | 2.4 | 1938 | 6 | 5169835-1 | Patent No. 5169835 |
| C 153 | 14 | 2.4 | 1113 | 2 | US-08-907-492A-3 | Sequence 3, App11 | 225 | 14 | 2.4 | 1976 | 3 | US-09-111-924-3 | Sequence 3, App11 |
| C 154 | 14 | 2.4 | 1200 | 3 | US-08-859-167-5 | Sequence 5, App11 | 226 | 14 | 2.4 | 2000 | 3 | US-09-165-042-2 | Sequence 2, App11 |
| C 155 | 14 | 2.4 | 1200 | 3 | US-09-109-273-5 | Sequence 5, App11 | 227 | 14 | 2.4 | 2000 | 3 | US-08-483-170-1 | Sequence 1, App11 |
| C 156 | 14 | 2.4 | 1200 | 4 | US-09-276-993-5 | Sequence 5, App11 | 228 | 14 | 2.4 | 2002 | 1 | US-08-867-331-1 | Sequence 1, App11 |
| C 157 | 14 | 2.4 | 1205 | 2 | US-09-154-802-2 | Sequence 2, App11 | 229 | 14 | 2.4 | 2064 | 1 | US-07-973-324A-1 | Sequence 1, App11 |
| C 158 | 14 | 2.4 | 1205 | 3 | US-09-373-029-2 | Sequence 2, App11 | 230 | 14 | 2.4 | 2086 | 1 | US-08-343-380-1 | Sequence 1, App11 |
| C 159 | 14 | 2.4 | 1268 | 1 | US-08-234-939-6 | Sequence 6, App11 | 231 | 14 | 2.4 | 2086 | 4 | US-09-072-435-1 | Sequence 1, App11 |
| C 160 | 14 | 2.4 | 1268 | 1 | US-08-558-865-6 | Sequence 6, App11 | 232 | 14 | 2.4 | 2115 | 5 | US-08-796-101-48 | Sequence 48, App1 |
| C 161 | 14 | 2.4 | 1268 | 1 | US-08-910-973-12 | Sequence 12, App1 | 233 | 14 | 2.4 | 2124 | 6 | 5428012-1 | Patent No. 5428012 |
| C 162 | 14 | 2.4 | 1269 | 5 | PCT-US93-00893-4 | Sequence 12, App1 | 234 | 14 | 2.4 | 2124 | 6 | 5451506-1 | Patent No. 5451506 |
| C 163 | 14 | 2.4 | 1289 | 2 | US-08-344-833-1 | Sequence 1, App11 | 235 | 14 | 2.4 | 2133 | 4 | US-09-187-124-1 | Sequence 1, App11 |
| C 164 | 14 | 2.4 | 1290 | 6 | 5182195-14 | Patent No. 5182195 | 236 | 14 | 2.4 | 2229 | 3 | US-08-888-077A-18 | Sequence 18, App1 |
| C 165 | 14 | 2.4 | 1317 | 3 | US-08-850-227-1 | Sequence 1, App11 | 237 | 14 | 2.4 | 2243 | 3 | US-09-084-079-4 | Sequence 4, App11 |
| C 166 | 14 | 2.4 | 1317 | 3 | US-08-850-227-5 | Sequence 5, App11 | 238 | 14 | 2.4 | 2276 | 2 | US-08-875-972-28 | Sequence 28, App1 |
| C 167 | 14 | 2.4 | 1317 | 3 | US-09-054-985A-1 | Sequence 5, App11 | 239 | 14 | 2.4 | 2285 | 2 | US-08-967-101-136 | Sequence 136, App |
| C 168 | 14 | 2.4 | 1317 | 3 | US-09-054-985A-5 | Sequence 5, App11 | 240 | 14 | 2.4 | 2285 | 2 | US-08-592-541-136 | Sequence 136, App |
| C 169 | 14 | 2.4 | 1353 | 4 | US-09-011-197-3 | Sequence 3, App11 | 241 | 14 | 2.4 | 2285 | 3 | US-09-124-698-136 | Sequence 136, App |
| C 170 | 14 | 2.4 | 1380 | 2 | US-08-467-559B-1 | Sequence 1, App11 | 242 | 14 | 2.4 | 2285 | 4 | US-09-127-480-136 | Sequence 136, App |
| C 171 | 14 | 2.4 | 1385 | 3 | US-08-935-263-3 | Sequence 3, App11 | 243 | 14 | 2.4 | 2296 | 4 | US-08-496-841C-137 | Sequence 137, App |
| C 172 | 14 | 2.4 | 1385 | 3 | US-08-234-939-5 | Sequence 3, App11 | 244 | 14 | 2.4 | 2335 | 2 | US-08-300-584-3 | Sequence 3, App11 |
| C 173 | 14 | 2.4 | 1385 | 1 | US-08-558-865-5 | Sequence 5, App11 | 245 | 14 | 2.4 | 2336 | 4 | US-08-476-123-3 | Sequence 3, App11 |
| C 173 | 14 | 2.4 | 1385 | 1 | US-08-558-865-5 | Sequence 5, App11 | 246 | 14 | 2.4 | 2336 | 4 | US-08-796-101-47 | Sequence 47, App1 |

| | | | | | | | | | | | | |
|-------|----|-----|------|---|-------------------|--------------------|-------|----|-------|---|-------------------|--------------------|
| C 247 | 14 | 2.4 | 2361 | 4 | US-09-011-197-1 | Sequence 1, Appl | 320 | 14 | 7032 | 3 | US-08-450-562-24 | Sequence 24, Appl |
| C 248 | 14 | 2.4 | 2409 | 1 | US-08-392-828C-1 | Sequence 1, Appl | 321 | 14 | 7089 | 3 | US-08-949-386-25 | Sequence 25, Appl |
| C 249 | 14 | 2.4 | 2409 | 3 | US-09-330-945-1 | Sequence 1, Appl | 322 | 14 | 7089 | 3 | US-08-450-562-25 | Sequence 25, Appl |
| C 250 | 14 | 2.4 | 2456 | 3 | US-08-813-150-5 | Sequence 5, Appl | 323 | 14 | 9045 | 3 | US-09-121-321-1 | Sequence 1, Appl |
| C 251 | 14 | 2.4 | 2529 | 3 | US-09-051-669A-5 | Sequence 5, Appl | 324 | 14 | 9045 | 3 | US-08-933-803A-1 | Sequence 1, Appl |
| C 252 | 14 | 2.4 | 2609 | 4 | US-09-141-212-7 | Sequence 7, Appl | 325 | 14 | 9370 | 4 | US-08-320-559-27 | Sequence 27, Appl |
| C 253 | 14 | 2.4 | 2625 | 4 | US-09-378-255-1 | Sequence 1, Appl | 326 | 14 | 9370 | 5 | US-08-545-860D-27 | Sequence 27, Appl |
| C 254 | 14 | 2.4 | 2669 | 4 | US-09-141-212-9 | Sequence 9, Appl | 327 | 14 | 9370 | 5 | PCT-US94-04496-27 | Sequence 27, Appl |
| C 255 | 14 | 2.4 | 2693 | 4 | US-09-141-212-5 | Sequence 5, Appl | 328 | 14 | 9391 | 1 | US-08-320-559-25 | Sequence 25, Appl |
| C 256 | 14 | 2.4 | 2709 | 4 | US-09-251-372-1 | Sequence 1, Appl | 329 | 14 | 9391 | 5 | US-08-545-860D-25 | Sequence 25, Appl |
| C 257 | 14 | 2.4 | 2733 | 1 | US-08-310-271-1 | Sequence 1, Appl | 330 | 14 | 9391 | 5 | PCT-US94-04496-25 | Sequence 25, Appl |
| C 258 | 14 | 2.4 | 2733 | 1 | US-09-032-742-9 | Sequence 9, Appl | C 331 | 14 | 9515 | 1 | US-08-920-812-13 | Sequence 13, Appl |
| C 259 | 14 | 2.4 | 2764 | 2 | US-08-465-971B-1 | Sequence 1, Appl | C 332 | 14 | 9515 | 1 | US-08-920-827-13 | Sequence 13, Appl |
| C 260 | 14 | 2.4 | 2816 | 2 | US-08-785-241-1 | Sequence 1, Appl | C 333 | 14 | 9515 | 1 | US-08-921-177-13 | Sequence 13, Appl |
| C 261 | 14 | 2.4 | 2943 | 2 | US-08-788-892-1 | Sequence 1, Appl | C 334 | 14 | 9515 | 1 | US-08-362-577C-13 | Sequence 13, Appl |
| C 262 | 14 | 2.4 | 2982 | 1 | US-08-348-143-2 | Sequence 2, Appl | C 335 | 14 | 9515 | 2 | US-08-920-828-13 | Sequence 13, Appl |
| C 263 | 14 | 2.4 | 2982 | 1 | US-08-571-785-3 | Sequence 2, Appl | C 336 | 14 | 10095 | 2 | US-08-822-586-45 | Sequence 45, Appl |
| C 264 | 14 | 2.4 | 2984 | 5 | PCT-US93-00893-1 | Sequence 1, Appl | C 337 | 14 | 10095 | 3 | US-08-822-586-45 | Sequence 45, Appl |
| C 265 | 14 | 2.4 | 2984 | 5 | PCT-US93-00893-2 | Sequence 2, Appl | C 338 | 14 | 10868 | 2 | US-09-058-411-1 | Sequence 1, Appl |
| C 266 | 14 | 2.4 | 2999 | 2 | US-09-014-569-8 | Sequence 8, Appl | 339 | 14 | 10898 | 3 | US-08-481-658B-5 | Sequence 5, Appl |
| C 267 | 14 | 2.4 | 3006 | 3 | US-09-552-351-1 | Sequence 1, Appl | 340 | 14 | 10898 | 2 | US-08-477-504A-5 | Sequence 5, Appl |
| C 268 | 14 | 2.4 | 3018 | 3 | US-08-942-572-1 | Sequence 1, Appl | 341 | 14 | 10898 | 2 | US-08-486-756A-5 | Sequence 5, Appl |
| C 269 | 14 | 2.4 | 3040 | 4 | US-09-378-255-3 | Sequence 3, Appl | 342 | 14 | 10898 | 2 | US-08-485-862B-5 | Sequence 5, Appl |
| C 270 | 14 | 2.4 | 3046 | 1 | US-08-726-725-1 | Sequence 1, Appl | 343 | 14 | 10898 | 3 | US-08-787-739-5 | Sequence 5, Appl |
| C 271 | 14 | 2.4 | 3240 | 4 | US-09-262-773-7 | Sequence 7, Appl | 344 | 14 | 10898 | 3 | US-08-487-077A-5 | Sequence 5, Appl |
| C 272 | 14 | 2.4 | 3244 | 4 | US-09-262-773-3 | Sequence 3, Appl | 345 | 14 | 10898 | 3 | US-08-485-863A-5 | Sequence 5, Appl |
| C 273 | 14 | 2.4 | 3255 | 2 | US-08-916-917-11 | Sequence 11, Appl | 346 | 14 | 10898 | 4 | US-08-485-049D-5 | Sequence 5, Appl |
| C 274 | 14 | 2.4 | 3255 | 3 | US-09-225-170-11 | Sequence 11, Appl | C 347 | 14 | 13417 | 2 | US-08-637-759B-37 | Sequence 37, Appl |
| C 275 | 14 | 2.4 | 3264 | 4 | US-09-262-773-5 | Sequence 5, Appl | C 348 | 14 | 13417 | 3 | US-08-871-353A-37 | Sequence 37, Appl |
| C 276 | 14 | 2.4 | 3268 | 4 | US-09-262-773-1 | Sequence 1, Appl | C 349 | 14 | 20137 | 4 | US-09-262-773-206 | Sequence 206, Appl |
| C 277 | 14 | 2.4 | 3268 | 4 | US-09-262-773-1 | Sequence 1, Appl | C 350 | 14 | 20138 | 4 | US-09-262-773-9 | Sequence 9, Appl |
| C 278 | 14 | 2.4 | 3300 | 2 | US-08-928-692-29 | Sequence 29, Appl | 351 | 14 | 20303 | 4 | US-08-370-975B-6 | Sequence 6, Appl |
| C 279 | 14 | 2.4 | 3426 | 1 | US-08-234-939-1 | Sequence 1, Appl | C 352 | 14 | 20710 | 1 | US-08-420-235B-1 | Sequence 1, Appl |
| C 280 | 14 | 2.4 | 3426 | 3 | US-08-654-025-6 | Sequence 1, Appl | C 353 | 14 | 20710 | 5 | US-08-793-624-1 | Sequence 1, Appl |
| C 281 | 14 | 2.4 | 3622 | 1 | US-08-306-691B-13 | Sequence 13, Appl | C 354 | 14 | 20710 | 5 | PCT-US93-10194-1 | Sequence 1, Appl |
| C 282 | 14 | 2.4 | 3622 | 5 | PCT-US93-06251-24 | Sequence 24, Appl | C 355 | 14 | 22071 | 4 | US-09-262-773-210 | Sequence 210, Appl |
| C 283 | 14 | 2.4 | 3627 | 1 | US-08-104-072B-6 | Sequence 6, Appl | 356 | 14 | 26700 | 1 | US-08-472-217-1 | Sequence 1, Appl |
| C 284 | 14 | 2.4 | 3627 | 1 | US-08-351-413-7 | Sequence 7, Appl | 357 | 14 | 26700 | 2 | US-08-488-199-5 | Sequence 5, Appl |
| C 285 | 14 | 2.4 | 3627 | 2 | US-09-025-583-7 | Sequence 7, Appl | 358 | 14 | 26700 | 3 | US-08-760-534A-1 | Sequence 1, Appl |
| C 286 | 14 | 2.4 | 3630 | 4 | US-09-378-255-5 | Sequence 5, Appl | 359 | 14 | 26764 | 1 | US-08-370-975B-1 | Sequence 1, Appl |
| C 287 | 14 | 2.4 | 3630 | 4 | US-09-141-212-1 | Sequence 1, Appl | 360 | 14 | 28804 | 2 | US-08-532-874-1 | Sequence 1, Appl |
| C 288 | 14 | 2.4 | 3630 | 4 | US-09-141-212-3 | Sequence 3, Appl | 361 | 14 | 28804 | 3 | US-09-096-867-2 | Sequence 2, Appl |
| C 289 | 14 | 2.4 | 3630 | 4 | US-09-552-351-3 | Sequence 3, Appl | 362 | 14 | 28804 | 3 | US-09-096-867-2 | Sequence 2, Appl |
| C 290 | 14 | 2.4 | 3630 | 4 | US-09-251-372-5 | Sequence 5, Appl | 363 | 14 | 35100 | 2 | US-08-770-379-18 | Sequence 18, Appl |
| C 291 | 14 | 2.4 | 3701 | 4 | US-08-845-258-10 | Sequence 10, Appl | C 364 | 14 | 35100 | 2 | US-08-770-379-19 | Sequence 19, Appl |
| C 292 | 14 | 2.4 | 3701 | 4 | US-08-990-571-10 | Sequence 10, Appl | C 365 | 14 | 35100 | 4 | US-08-757-669A-18 | Sequence 18, Appl |
| C 293 | 14 | 2.4 | 3744 | 2 | US-08-348-353-16 | Sequence 16, Appl | C 366 | 14 | 35100 | 4 | US-08-757-669A-19 | Sequence 19, Appl |
| C 294 | 14 | 2.4 | 3744 | 2 | US-08-465-965-16 | Sequence 16, Appl | C 367 | 14 | 49272 | 1 | US-08-614-770A-1 | Sequence 1, Appl |
| C 295 | 14 | 2.4 | 3744 | 3 | US-08-465-966-16 | Sequence 16, Appl | C 368 | 14 | 80161 | 3 | US-09-036-987A-1 | Sequence 1, Appl |
| C 296 | 14 | 2.4 | 4027 | 1 | US-08-348-143-3 | Sequence 3, Appl | 369 | 14 | 87350 | 3 | US-08-781-891-79 | Sequence 79, Appl |
| C 297 | 14 | 2.4 | 4027 | 1 | US-08-348-143-4 | Sequence 4, Appl | C 370 | 14 | 87350 | 3 | US-08-882-046-81 | Sequence 81, Appl |
| C 298 | 14 | 2.4 | 4027 | 1 | US-08-571-785-3 | Sequence 3, Appl | C 371 | 14 | 19 | 1 | US-08-410-540-5 | Sequence 5, Appl |
| C 299 | 14 | 2.4 | 4027 | 1 | US-08-571-785-4 | Sequence 4, Appl | C 372 | 14 | 21 | 1 | US-08-372-892-5 | Sequence 5, Appl |
| C 300 | 14 | 2.4 | 4190 | 4 | US-08-924-345-1 | Sequence 1, Appl | C 373 | 14 | 21 | 1 | US-08-357-565-31 | Sequence 31, Appl |
| C 301 | 14 | 2.4 | 4792 | 3 | US-08-781-891-205 | Sequence 205, Appl | C 374 | 14 | 21 | 1 | US-08-252-626A-4 | Sequence 4, Appl |
| C 302 | 14 | 2.4 | 4922 | 2 | US-08-330-272-5 | Sequence 5, Appl | C 375 | 14 | 21 | 1 | US-08-454-720A-31 | Sequence 31, Appl |
| C 303 | 14 | 2.4 | 4922 | 5 | PCT-US95-13663-5 | Sequence 5, Appl | C 376 | 14 | 21 | 3 | US-08-987-326-21 | Sequence 21, Appl |
| C 304 | 14 | 2.4 | 5183 | 3 | US-09-039-555B-18 | Sequence 18, Appl | C 377 | 14 | 21 | 3 | US-08-987-326-21 | Sequence 21, Appl |
| C 305 | 14 | 2.4 | 5275 | 4 | US-08-796-101-49 | Sequence 49, Appl | C 378 | 14 | 21 | 3 | US-08-353-901-6 | Sequence 6, Appl |
| C 306 | 14 | 2.4 | 5275 | 4 | US-08-796-101-49 | Sequence 49, Appl | C 379 | 14 | 21 | 4 | US-09-436-605-17 | Sequence 17, Appl |
| C 307 | 14 | 2.4 | 5470 | 4 | US-08-441-139-12 | Sequence 12, Appl | C 380 | 14 | 21 | 6 | US-09-195-817-6 | Sequence 6, Appl |
| C 308 | 14 | 2.4 | 5470 | 6 | 5196523-5 | Patent No. 5196523 | C 381 | 14 | 27 | 1 | US-08-221-817-6 | Sequence 6, Appl |
| C 309 | 14 | 2.4 | 5529 | 3 | US-08-592-214A-20 | Sequence 20, Appl | C 382 | 14 | 27 | 1 | US-08-454-439-6 | Sequence 6, Appl |
| C 310 | 14 | 2.4 | 5855 | 3 | US-09-149-976-20 | Sequence 20, Appl | C 383 | 14 | 27 | 5 | PCT-US94-10487-6 | Sequence 6, Appl |
| C 311 | 14 | 2.4 | 5855 | 3 | US-09-149-976-20 | Sequence 20, Appl | C 384 | 14 | 27 | 5 | US-08-151-331A-7 | Sequence 7, Appl |
| C 312 | 14 | 2.4 | 6317 | 1 | US-08-920-812-21 | Sequence 21, Appl | C 385 | 14 | 38 | 1 | US-08-453-588-5 | Sequence 5, Appl |
| C 313 | 14 | 2.4 | 6317 | 1 | US-08-920-827-21 | Sequence 21, Appl | C 386 | 14 | 39 | 1 | US-08-644-664B-4 | Sequence 4, Appl |
| C 314 | 14 | 2.4 | 6317 | 1 | US-08-921-177-21 | Sequence 21, Appl | C 387 | 14 | 39 | 1 | US-08-644-664B-5 | Sequence 5, Appl |
| C 315 | 14 | 2.4 | 6317 | 1 | US-08-362-577C-21 | Sequence 21, Appl | C 388 | 14 | 39 | 2 | US-08-761-277A-4 | Sequence 4, Appl |
| C 316 | 14 | 2.4 | 6317 | 2 | US-08-362-577C-21 | Sequence 21, Appl | C 389 | 14 | 39 | 2 | US-08-761-277A-5 | Sequence 5, Appl |
| C 317 | 14 | 2.4 | 6476 | 4 | US-08-920-828-21 | Sequence 21, Appl | C 390 | 14 | 39 | 3 | US-08-521-079-5 | Sequence 5, Appl |
| C 318 | 14 | 2.4 | 7032 | 4 | US-08-149-097D-24 | Sequence 24, Appl | C 391 | 14 | 45 | 1 | US-08-453-588-29 | Sequence 29, Appl |
| C 319 | 14 | 2.4 | 7032 | 3 | US-08-949-386-24 | Sequence 24, Appl | C 392 | 14 | 45 | 3 | US-08-521-079-29 | Sequence 29, Appl |

| | | | | | | |
|-----|----|-----|-----|---|--------------------|--------------------|
| 393 | 13 | 2.2 | 51 | 1 | US-08-126-564A-37 | Sequence 37, Appl |
| 394 | 13 | 2.2 | 51 | 1 | US-08-453-588-13 | Sequence 13, Appl |
| 395 | 13 | 2.2 | 51 | 2 | US-08-850-049-37 | Sequence 37, Appl |
| 396 | 13 | 2.2 | 51 | 2 | US-08-050-478-37 | Sequence 37, Appl |
| 397 | 13 | 2.2 | 51 | 3 | US-08-521-079-13 | Sequence 13, Appl |
| 398 | 13 | 2.2 | 51 | 3 | PCT-US94-09143-37 | Sequence 37, Appl |
| 399 | 13 | 2.2 | 69 | 2 | US-08-180-524-4 | Sequence 4, Appl1 |
| 400 | 13 | 2.2 | 69 | 2 | US-08-975-166-4 | Sequence 4, Appl1 |
| 401 | 13 | 2.2 | 80 | 2 | US-07-733-095B-10 | Sequence 10, Appl |
| 402 | 13 | 2.2 | 110 | 1 | US-07-668-648-11 | Sequence 11, Appl |
| 403 | 13 | 2.2 | 110 | 2 | US-08-429-998-11 | Sequence 11, Appl |
| 404 | 13 | 2.2 | 110 | 2 | US-08-431-333-11 | Sequence 11, Appl |
| 405 | 13 | 2.2 | 110 | 5 | PCT-US91-02321-11 | Sequence 11, Appl |
| 406 | 13 | 2.2 | 120 | 1 | US-07-666-719-21 | Sequence 21, Appl |
| 407 | 13 | 2.2 | 120 | 6 | 5191064-1 | Patent No. 5191064 |
| 408 | 13 | 2.2 | 134 | 1 | US-07-998-003A-78 | Sequence 78, Appl |
| 409 | 13 | 2.2 | 134 | 1 | US-08-453-274B-78 | Sequence 78, Appl |
| 410 | 13 | 2.2 | 134 | 1 | US-08-453-695A-78 | Sequence 78, Appl |
| 411 | 13 | 2.2 | 134 | 1 | US-08-268-161A-78 | Sequence 78, Appl |
| 412 | 13 | 2.2 | 134 | 2 | US-08-453-702A-78 | Sequence 78, Appl |
| 413 | 13 | 2.2 | 134 | 5 | PCT-US93-12588-78 | Sequence 78, Appl |
| 414 | 13 | 2.2 | 134 | 5 | PCT-US95-08071-78 | Sequence 78, Appl |
| 415 | 13 | 2.2 | 142 | 1 | US-08-488-144-29 | Sequence 29, Appl |
| 416 | 13 | 2.2 | 146 | 2 | US-08-180-524-3 | Sequence 3, Appl1 |
| 417 | 13 | 2.2 | 146 | 2 | US-08-975-166-3 | Sequence 3, Appl1 |
| 418 | 13 | 2.2 | 155 | 2 | US-07-733-093B-12 | Sequence 12, Appl |
| 419 | 13 | 2.2 | 171 | 6 | 546585-4 | Patent No. 546585 |
| 420 | 13 | 2.2 | 180 | 6 | 5437990-1 | Patent No. 5437990 |
| 421 | 13 | 2.2 | 186 | 4 | US-08-905-223-43 | Sequence 43, Appl |
| 422 | 13 | 2.2 | 196 | 5 | PCT-US93-06251-15 | Sequence 15, Appl |
| 423 | 13 | 2.2 | 198 | 2 | US-08-588-258B-16 | Sequence 16, Appl |
| 424 | 13 | 2.2 | 198 | 3 | US-08-460-503-16 | Sequence 16, Appl |
| 425 | 13 | 2.2 | 198 | 5 | PCT-US96-08295-16 | Sequence 16, Appl |
| 426 | 13 | 2.2 | 207 | 1 | US-08-435-684A-40 | Sequence 40, Appl |
| 427 | 13 | 2.2 | 207 | 2 | US-08-934-877A-40 | Sequence 40, Appl |
| 428 | 13 | 2.2 | 207 | 4 | US-08-871-678C-40 | Sequence 40, Appl |
| 429 | 13 | 2.2 | 211 | 4 | US-08-905-223-213 | Sequence 213, App |
| 430 | 13 | 2.2 | 239 | 1 | US-08-337-268A-44 | Sequence 44, Appl |
| 431 | 13 | 2.2 | 239 | 1 | US-08-484-570A-44 | Sequence 44, Appl |
| 432 | 13 | 2.2 | 249 | 1 | US-08-729-447-19 | Sequence 19, Appl |
| 433 | 13 | 2.2 | 253 | 6 | 546585-5 | Patent No. 546585 |
| 434 | 13 | 2.2 | 254 | 4 | US-09-318-758-6 | Sequence 6, Appl1 |
| 435 | 13 | 2.2 | 285 | 4 | US-09-060-756-135 | Sequence 135, App |
| 436 | 13 | 2.2 | 292 | 2 | US-08-245-511-15 | Sequence 15, Appl |
| 437 | 13 | 2.2 | 292 | 2 | US-08-600-993A-15 | Sequence 15, Appl |
| 438 | 13 | 2.2 | 297 | 2 | US-08-468-819A-89 | Sequence 89, Appl |
| 439 | 13 | 2.2 | 329 | 1 | US-08-148-910-2 | Sequence 2, Appl1 |
| 440 | 13 | 2.2 | 329 | 1 | US-08-148-910-2 | Sequence 2, Appl1 |
| 441 | 13 | 2.2 | 329 | 1 | US-08-148-910-13 | Sequence 13, Appl |
| 442 | 13 | 2.2 | 329 | 1 | US-08-148-910-13 | Sequence 13, Appl |
| 443 | 13 | 2.2 | 329 | 1 | US-08-448-937A-2 | Sequence 2, Appl1 |
| 444 | 13 | 2.2 | 329 | 1 | US-08-448-937A-2 | Sequence 2, Appl1 |
| 445 | 13 | 2.2 | 329 | 1 | US-08-448-937A-13 | Sequence 13, Appl |
| 446 | 13 | 2.2 | 329 | 1 | US-08-448-937A-13 | Sequence 13, Appl |
| 447 | 13 | 2.2 | 329 | 1 | US-08-358-171-20 | Sequence 20, Appl |
| 448 | 13 | 2.2 | 329 | 3 | US-09-090-947-20 | Sequence 20, Appl |
| 449 | 13 | 2.2 | 332 | 2 | US-08-469-412A-8 | Sequence 8, Appl1 |
| 450 | 13 | 2.2 | 332 | 4 | US-09-021-715-8 | Sequence 8, Appl1 |
| 451 | 13 | 2.2 | 335 | 4 | US-09-060-756-715 | Sequence 715, App |
| 452 | 13 | 2.2 | 342 | 2 | US-08-475-844-15 | Sequence 15, Appl |
| 453 | 13 | 2.2 | 342 | 3 | PCT-US95-08429-15 | Sequence 15, Appl |
| 454 | 13 | 2.2 | 348 | 3 | US-08-483-749A-17 | Sequence 17, Appl |
| 455 | 13 | 2.2 | 357 | 2 | US-08-318-157B-1 | Sequence 1, Appl1 |
| 456 | 13 | 2.2 | 363 | 3 | US-08-599-252-100 | Sequence 100, App |
| 457 | 13 | 2.2 | 363 | 5 | PCT-US96-06352-100 | Sequence 100, App |
| 458 | 13 | 2.2 | 363 | 5 | PCT-US96-06352-100 | Sequence 100, App |
| 459 | 13 | 2.2 | 368 | 2 | US-08-843-530B-21 | Sequence 21, Appl |
| 460 | 13 | 2.2 | 371 | 1 | US-08-335-573-5 | Sequence 5, Appl1 |
| 461 | 13 | 2.2 | 371 | 2 | US-08-522-841-6 | Sequence 6, Appl1 |
| 462 | 13 | 2.2 | 403 | 4 | US-09-060-756-391 | Sequence 391, App |
| 463 | 13 | 2.2 | 417 | 4 | US-09-136-315-1 | Sequence 1, Appl1 |
| 464 | 13 | 2.2 | 417 | 4 | US-09-136-315-5 | Sequence 5, Appl1 |
| 465 | 13 | 2.2 | 420 | 1 | US-08-149-099C-9 | Sequence 9, Appl1 |

| | | | | | | |
|-----|----|-----|-----|---|---------------------|--------------------|
| 466 | 13 | 2.2 | 420 | 1 | US-08-476-275-5 | Sequence 5, Appl1 |
| 467 | 13 | 2.2 | 420 | 2 | US-08-478-967A-9 | Sequence 9, Appl1 |
| 468 | 13 | 2.2 | 420 | 2 | US-09-120-663-3 | Sequence 3, Appl1 |
| 469 | 13 | 2.2 | 429 | 4 | US-08-905-223-61 | Sequence 61, Appl |
| 470 | 13 | 2.2 | 442 | 4 | US-09-060-756-293 | Sequence 293, App |
| 471 | 13 | 2.2 | 444 | 3 | US-08-873-709-1 | Sequence 1, Appl1 |
| 472 | 13 | 2.2 | 456 | 1 | US-08-450-945-70 | Sequence 70, Appl |
| 473 | 13 | 2.2 | 456 | 4 | US-08-976-161-70 | Sequence 70, Appl |
| 474 | 13 | 2.2 | 457 | 4 | US-09-060-756-591 | Sequence 591, App |
| 475 | 13 | 2.2 | 511 | 3 | US-09-188-930-216 | Sequence 216, App |
| 476 | 13 | 2.2 | 534 | 2 | US-09-098-487-11 | Sequence 11, Appl |
| 477 | 13 | 2.2 | 545 | 4 | US-08-446-935-1 | Sequence 5, Appl1 |
| 478 | 13 | 2.2 | 554 | 1 | US-08-330-123A-1 | Sequence 1, Appl1 |
| 479 | 13 | 2.2 | 561 | 2 | US-08-655-821-1 | Sequence 1, Appl1 |
| 480 | 13 | 2.2 | 578 | 4 | US-08-991-789A-222 | Sequence 222, App |
| 481 | 13 | 2.2 | 609 | 1 | US-08-087-772A-6 | Sequence 6, Appl1 |
| 482 | 13 | 2.2 | 620 | 2 | US-08-522-841-4 | Sequence 4, Appl1 |
| 483 | 13 | 2.2 | 621 | 2 | US-08-522-841-5 | Sequence 5, Appl1 |
| 484 | 13 | 2.2 | 627 | 1 | US-08-284-784-26 | Sequence 26, Appl |
| 485 | 13 | 2.2 | 627 | 1 | US-08-854-811-26 | Sequence 26, Appl |
| 486 | 13 | 2.2 | 641 | 6 | 546585-6 | Patent No. 546585 |
| 487 | 13 | 2.2 | 660 | 4 | US-08-990-791-1 | Sequence 1, Appl1 |
| 488 | 13 | 2.2 | 681 | 6 | 5187078-24 | Patent No. 5187078 |
| 489 | 13 | 2.2 | 686 | 4 | US-08-898-416-715 | Sequence 715, App |
| 490 | 13 | 2.2 | 686 | 4 | US-08-898-416-715 | Sequence 715, App |
| 491 | 13 | 2.2 | 707 | 2 | US-08-850-910A-40 | Sequence 40, Appl |
| 492 | 13 | 2.2 | 707 | 2 | US-08-850-910A-40 | Sequence 40, Appl |
| 493 | 13 | 2.2 | 716 | 4 | US-08-998-416-1129 | Sequence 1129, Ap |
| 494 | 13 | 2.2 | 719 | 4 | US-09-411-578-39 | Sequence 39, Appl |
| 495 | 13 | 2.2 | 720 | 1 | US-08-284-784A-22 | Sequence 22, Appl |
| 496 | 13 | 2.2 | 720 | 2 | US-08-854-811-22 | Sequence 22, Appl |
| 497 | 13 | 2.2 | 732 | 4 | US-08-998-416-927 | Sequence 927, App |
| 498 | 13 | 2.2 | 745 | 1 | US-08-036-555B-163 | Sequence 163, App |
| 499 | 13 | 2.2 | 745 | 1 | US-08-469-569-163 | Sequence 163, App |
| 500 | 13 | 2.2 | 745 | 1 | US-08-249-322A-163 | Sequence 163, App |
| 501 | 13 | 2.2 | 745 | 1 | US-08-469-526A-163 | Sequence 163, App |
| 502 | 13 | 2.2 | 745 | 2 | US-08-734-591A-163 | Sequence 163, App |
| 503 | 13 | 2.2 | 745 | 2 | US-08-469-660-163 | Sequence 163, App |
| 504 | 13 | 2.2 | 745 | 3 | US-08-341-018-51 | Sequence 51, Appl |
| 505 | 13 | 2.2 | 745 | 3 | US-08-470-335-163 | Sequence 163, App |
| 506 | 13 | 2.2 | 745 | 4 | US-08-735-021-163 | Sequence 163, App |
| 507 | 13 | 2.2 | 745 | 4 | US-08-734-664A-163 | Sequence 163, App |
| 508 | 13 | 2.2 | 745 | 4 | US-08-470-339-163 | Sequence 163, App |
| 509 | 13 | 2.2 | 745 | 5 | PCT-US94-05083C-159 | Sequence 159, App |
| 510 | 13 | 2.2 | 745 | 5 | PCT-US95-06864A-163 | Sequence 163, App |
| 511 | 13 | 2.2 | 747 | 4 | US-08-998-416-246 | Sequence 246, App |
| 512 | 13 | 2.2 | 754 | 1 | US-07-958-140-1 | Sequence 1, Appl1 |
| 513 | 13 | 2.2 | 754 | 5 | PCT-US93-09166-1 | Sequence 1, Appl1 |
| 514 | 13 | 2.2 | 758 | 4 | US-08-998-416-435 | Sequence 435, App |
| 515 | 13 | 2.2 | 799 | 4 | US-08-998-416-420 | Sequence 420, App |
| 516 | 13 | 2.2 | 801 | 2 | US-08-770-379-16 | Sequence 16, Appl |
| 517 | 13 | 2.2 | 801 | 4 | US-08-757-669A-16 | Sequence 16, Appl |
| 518 | 13 | 2.2 | 807 | 1 | US-08-566-369-1 | Sequence 1, Appl1 |
| 519 | 13 | 2.2 | 807 | 1 | US-08-566-369-1 | Sequence 1, Appl1 |
| 520 | 13 | 2.2 | 807 | 4 | US-09-074-331-1 | Sequence 4, Appl1 |
| 521 | 13 | 2.2 | 807 | 4 | US-09-074-331-4 | Sequence 4, Appl1 |
| 522 | 13 | 2.2 | 807 | 5 | PCT-US95-01937-1 | Sequence 1, Appl1 |
| 523 | 13 | 2.2 | 807 | 5 | PCT-US95-01937-4 | Sequence 4, Appl1 |
| 524 | 13 | 2.2 | 812 | 4 | US-08-998-416-410 | Sequence 410, App |
| 525 | 13 | 2.2 | 817 | 4 | US-08-998-416-374 | Sequence 374, App |
| 526 | 13 | 2.2 | 832 | 1 | US-08-340-539A-6 | Sequence 6, Appl1 |
| 527 | 13 | 2.2 | 832 | 2 | US-08-461-592B-6 | Sequence 6, Appl1 |
| 528 | 13 | 2.2 | 836 | 4 | US-08-971-090-6 | Sequence 6, Appl1 |
| 529 | 13 | 2.2 | 836 | 4 | US-08-971-090-7 | Sequence 7, Appl1 |
| 530 | 13 | 2.2 | 836 | 4 | US-09-187-946-5 | Sequence 5, Appl1 |
| 531 | 13 | 2.2 | 838 | 5 | PCT-US94-10457-1 | Sequence 7, Appl1 |
| 532 | 13 | 2.2 | 838 | 5 | PCT-US96-03965-7 | Sequence 7, Appl1 |
| 533 | 13 | 2.2 | 849 | 3 | US-08-807-300-2 | Sequence 2, Appl1 |
| 534 | 13 | 2.2 | 853 | 3 | US-08-963-409-2 | Sequence 2, Appl1 |
| 535 | 13 | 2.2 | 854 | 3 | US-09-064-693A-24 | Sequence 24, Appl |
| 536 | 13 | 2.2 | 861 | 4 | US-08-953-326-8 | Sequence 8, Appl1 |
| 537 | 13 | 2.2 | 863 | 4 | US-08-998-416-489 | Sequence 489, App |
| 538 | 13 | 2.2 | 867 | 3 | US-08-869-696-6 | Sequence 6, Appl1 |

| | | | | | | | | | | | | | |
|-----|----|-----|------|---|-------------------|-------------------|-------|----|-----|------|---|-------------------|-------------------|
| 539 | 13 | 2.2 | 867 | 4 | US-09-475-316A-22 | Sequence 22, Appl | C 612 | 13 | 2.2 | 1240 | 2 | US-08-748-428-1 | Sequence 1, Appl |
| 540 | 13 | 2.2 | 870 | 2 | US-08-403-853-15 | Sequence 15, Appl | C 613 | 13 | 2.2 | 1268 | 1 | US-08-910-973-12 | Sequence 12, Appl |
| 541 | 13 | 2.2 | 903 | 1 | US-08-944-604-15 | Sequence 15, Appl | 614 | 13 | 2.2 | 1269 | 1 | US-08-396-218-1 | Sequence 1, Appl |
| 542 | 13 | 2.2 | 921 | 1 | US-08-420-235B-30 | Sequence 30, Appl | 615 | 13 | 2.2 | 1269 | 1 | US-08-760-116-1 | Sequence 1, Appl |
| 543 | 13 | 2.2 | 921 | 3 | US-08-793-624-30 | Sequence 30, Appl | 616 | 13 | 2.2 | 1282 | 3 | US-08-869-636-21 | Sequence 21, Appl |
| 544 | 13 | 2.2 | 921 | 3 | PCT-US95-10194-30 | Sequence 30, Appl | 617 | 13 | 2.2 | 1298 | 3 | US-08-948-705-3 | Sequence 3, Appl |
| 545 | 13 | 2.2 | 925 | 3 | US-08-858-003-1 | Sequence 1, Appl | 618 | 13 | 2.2 | 1299 | 1 | US-08-453-472-3 | Sequence 3, Appl |
| 546 | 13 | 2.2 | 925 | 3 | US-09-078-166-1 | Sequence 1, Appl | 619 | 13 | 2.2 | 1299 | 1 | US-08-038-948-3 | Sequence 3, Appl |
| 547 | 13 | 2.2 | 925 | 4 | US-08-997-467-1 | Sequence 1, Appl | 620 | 13 | 2.2 | 1299 | 1 | US-08-453-952-3 | Sequence 3, Appl |
| 548 | 13 | 2.2 | 946 | 2 | US-08-816-605-1 | Sequence 1, Appl | 621 | 13 | 2.2 | 1299 | 2 | US-08-862-903-3 | Sequence 3, Appl |
| 549 | 13 | 2.2 | 965 | 1 | US-08-736-361A-3 | Sequence 3, Appl | 622 | 13 | 2.2 | 1299 | 2 | US-08-484-158B-60 | Sequence 60, Appl |
| 550 | 13 | 2.2 | 967 | 3 | US-09-248-335-51 | Sequence 51, Appl | C 623 | 13 | 2.2 | 1312 | 5 | US-08-205-506A-1 | Sequence 1, Appl |
| 551 | 13 | 2.2 | 970 | 1 | US-08-148-910-3 | Sequence 3, Appl | C 624 | 13 | 2.2 | 1312 | 5 | PCT-US94-02389-1 | Sequence 1, Appl |
| 552 | 13 | 2.2 | 970 | 1 | US-08-148-910-3 | Sequence 3, Appl | 625 | 13 | 2.2 | 1325 | 2 | US-08-484-993B-11 | Sequence 11, Appl |
| 553 | 13 | 2.2 | 970 | 1 | US-08-448-937A-3 | Sequence 3, Appl | 626 | 13 | 2.2 | 1325 | 2 | US-08-484-956A-11 | Sequence 11, Appl |
| 554 | 13 | 2.2 | 970 | 1 | US-08-448-937A-3 | Sequence 3, Appl | 627 | 13 | 2.2 | 1325 | 2 | US-08-484-158B-11 | Sequence 11, Appl |
| 555 | 13 | 2.2 | 983 | 1 | US-08-886-765-1 | Sequence 1, Appl | 628 | 13 | 2.2 | 1325 | 2 | US-08-480-150A-11 | Sequence 11, Appl |
| 556 | 13 | 2.2 | 983 | 4 | US-09-115-660-1 | Sequence 1, Appl | 629 | 13 | 2.2 | 1325 | 3 | US-08-458-721-11 | Sequence 11, Appl |
| 557 | 13 | 2.2 | 996 | 1 | US-07-975-526-2 | Sequence 2, Appl | 630 | 13 | 2.2 | 1325 | 3 | US-08-149-223A-11 | Sequence 11, Appl |
| 558 | 13 | 2.2 | 996 | 5 | PCT-US95-15353-38 | Sequence 38, Appl | 631 | 13 | 2.2 | 1360 | 1 | US-08-087-772A-4 | Sequence 4, Appl |
| 559 | 13 | 2.2 | 996 | 5 | PCT-US95-15353-41 | Sequence 40, Appl | 632 | 13 | 2.2 | 1372 | 3 | US-08-869-666-22 | Sequence 22, Appl |
| 560 | 13 | 2.2 | 1000 | 4 | US-09-018-584A-41 | Sequence 41, Appl | 633 | 13 | 2.2 | 1377 | 1 | US-07-648-796A-13 | Sequence 13, Appl |
| 561 | 13 | 2.2 | 1005 | 2 | US-08-657-866-10 | Sequence 10, Appl | 634 | 13 | 2.2 | 1400 | 1 | US-08-085-122-12 | Sequence 12, Appl |
| 562 | 13 | 2.2 | 1007 | 2 | US-08-665-647-2 | Sequence 2, Appl | 635 | 13 | 2.2 | 1400 | 4 | US-08-442-108B-22 | Sequence 22, Appl |
| 563 | 13 | 2.2 | 1013 | 4 | US-09-084-120-17 | Sequence 17, Appl | C 636 | 13 | 2.2 | 1415 | 1 | US-08-236-918A-7 | Sequence 7, Appl |
| 564 | 13 | 2.2 | 1014 | 1 | US-07-648-796A-9 | Sequence 9, Appl | 637 | 13 | 2.2 | 1417 | 2 | US-08-713-828-2 | Sequence 2, Appl |
| 565 | 13 | 2.2 | 1018 | 1 | US-08-444-083-6 | Sequence 6, Appl | 638 | 13 | 2.2 | 1417 | 2 | US-08-919-627-2 | Sequence 2, Appl |
| 566 | 13 | 2.2 | 1018 | 1 | US-08-286-304-6 | Sequence 6, Appl | 639 | 13 | 2.2 | 1417 | 2 | US-09-096-245-2 | Sequence 2, Appl |
| 567 | 13 | 2.2 | 1018 | 1 | US-08-442-745-6 | Sequence 6, Appl | 640 | 13 | 2.2 | 1419 | 3 | US-09-354-129-1 | Sequence 3, Appl |
| 568 | 13 | 2.2 | 1018 | 1 | US-08-443-129-6 | Sequence 6, Appl | 641 | 13 | 2.2 | 1425 | 2 | US-08-578-551-2 | Sequence 2, Appl |
| 569 | 13 | 2.2 | 1018 | 1 | US-08-443-952-6 | Sequence 6, Appl | 642 | 13 | 2.2 | 1425 | 2 | US-09-190-982-2 | Sequence 2, Appl |
| 570 | 13 | 2.2 | 1018 | 1 | US-08-443-130-6 | Sequence 6, Appl | 643 | 13 | 2.2 | 1425 | 4 | US-09-408-257-2 | Sequence 2, Appl |
| 571 | 13 | 2.2 | 1018 | 3 | US-08-896-095-2 | Sequence 2, Appl | 644 | 13 | 2.2 | 1426 | 4 | US-09-193-377B-7 | Sequence 7, Appl |
| 572 | 13 | 2.2 | 1018 | 3 | US-08-898-911-6 | Sequence 6, Appl | 645 | 13 | 2.2 | 1433 | 2 | US-08-666-392A-1 | Sequence 1, Appl |
| 573 | 13 | 2.2 | 1018 | 5 | PCT-US95-04467-6 | Sequence 6, Appl | C 646 | 13 | 2.2 | 1443 | 2 | US-08-403-853-19 | Sequence 19, Appl |
| 574 | 13 | 2.2 | 1021 | 1 | US-09-095-117-5 | Sequence 5, Appl | C 647 | 13 | 2.2 | 1462 | 1 | US-08-358-782D-14 | Sequence 14, Appl |
| 575 | 13 | 2.2 | 1029 | 1 | US-07-809-457A-15 | Sequence 15, Appl | C 648 | 13 | 2.2 | 1462 | 2 | US-08-764-527A-14 | Sequence 14, Appl |
| 576 | 13 | 2.2 | 1029 | 1 | US-08-553-943-15 | Sequence 15, Appl | C 649 | 13 | 2.2 | 1483 | 3 | US-08-833-610-1 | Sequence 1, Appl |
| 577 | 13 | 2.2 | 1029 | 5 | PCT-US91-09437-15 | Sequence 15, Appl | C 650 | 13 | 2.2 | 1485 | 2 | US-08-759-581B-12 | Sequence 12, Appl |
| 578 | 13 | 2.2 | 1030 | 4 | US-09-095-117-7 | Sequence 7, Appl | 651 | 13 | 2.2 | 1486 | 1 | US-08-250-859-1 | Sequence 1, Appl |
| 579 | 13 | 2.2 | 1044 | 3 | US-08-869-666-20 | Sequence 20, Appl | 652 | 13 | 2.2 | 1486 | 1 | US-08-347-029-1 | Sequence 1, Appl |
| 580 | 13 | 2.2 | 1045 | 3 | US-08-859-167-3 | Sequence 3, Appl | 653 | 13 | 2.2 | 1486 | 1 | US-08-490-803-1 | Sequence 1, Appl |
| 581 | 13 | 2.2 | 1045 | 3 | US-09-109-273-3 | Sequence 3, Appl | 654 | 13 | 2.2 | 1486 | 2 | US-08-457-254-1 | Sequence 1, Appl |
| 582 | 13 | 2.2 | 1045 | 4 | US-09-276-993-3 | Sequence 3, Appl | 655 | 13 | 2.2 | 1486 | 2 | US-08-484-257-1 | Sequence 1, Appl |
| 583 | 13 | 2.2 | 1048 | 4 | US-08-195-967-3 | Sequence 3, Appl | 656 | 13 | 2.2 | 1486 | 3 | US-08-999-927-3 | Sequence 3, Appl |
| 584 | 13 | 2.2 | 1057 | 1 | US-08-470-202-44 | Sequence 44, Appl | 657 | 13 | 2.2 | 1486 | 5 | PCT-US94-08806-1 | Sequence 1, Appl |
| 585 | 13 | 2.2 | 1057 | 1 | US-08-471-770-44 | Sequence 44, Appl | 658 | 13 | 2.2 | 1486 | 5 | PCT-US95-01175-1 | Sequence 1, Appl |
| 586 | 13 | 2.2 | 1057 | 2 | US-08-426-169-44 | Sequence 44, Appl | 659 | 13 | 2.2 | 1486 | 5 | PCT-US95-01829-3 | Sequence 3, Appl |
| 587 | 13 | 2.2 | 1061 | 1 | US-08-426-169-44 | Sequence 4, Appl | 660 | 13 | 2.2 | 1486 | 5 | PCT-US95-14932-1 | Sequence 1, Appl |
| 588 | 13 | 2.2 | 1061 | 4 | US-09-233-813-4 | Sequence 4, Appl | 661 | 13 | 2.2 | 1486 | 5 | PCT-US95-16626-1 | Sequence 1, Appl |
| 589 | 13 | 2.2 | 1061 | 5 | PCT-US95-09470-4 | Sequence 4, Appl | 662 | 13 | 2.2 | 1494 | 3 | US-08-724-466B-5 | Sequence 5, Appl |
| 590 | 13 | 2.2 | 1062 | 1 | US-07-603-133B-21 | Sequence 21, Appl | 663 | 13 | 2.2 | 1504 | 2 | US-08-850-910A-17 | Sequence 17, Appl |
| 591 | 13 | 2.2 | 1089 | 1 | US-08-154-915-1 | Sequence 1, Appl | C 664 | 13 | 2.2 | 1504 | 2 | US-08-850-910A-17 | Sequence 17, Appl |
| 592 | 13 | 2.2 | 1089 | 2 | US-08-464-517-37 | Sequence 37, Appl | 665 | 13 | 2.2 | 1506 | 2 | US-08-206-790A-22 | Sequence 22, Appl |
| 593 | 13 | 2.2 | 1089 | 2 | US-08-246-361A-37 | Sequence 37, Appl | 666 | 13 | 2.2 | 1506 | 5 | PCT-US95-02943-22 | Sequence 22, Appl |
| 594 | 13 | 2.2 | 1089 | 3 | US-08-463-772-37 | Sequence 37, Appl | 667 | 13 | 2.2 | 1507 | 2 | US-08-850-910A-38 | Sequence 38, Appl |
| 595 | 13 | 2.2 | 1089 | 5 | PCT-US93-09945-10 | Sequence 10, Appl | C 668 | 13 | 2.2 | 1507 | 2 | US-08-850-910A-38 | Sequence 38, Appl |
| 596 | 13 | 2.2 | 1089 | 5 | PCT-US93-09945-1 | Sequence 1, Appl | C 669 | 13 | 2.2 | 1518 | 3 | US-08-820-170A-20 | Sequence 20, Appl |
| 597 | 13 | 2.2 | 1119 | 2 | US-08-626-685A-7 | Sequence 7, Appl | C 670 | 13 | 2.2 | 1518 | 3 | US-09-035-699-20 | Sequence 20, Appl |
| 598 | 13 | 2.2 | 1132 | 3 | US-08-651-136C-21 | Sequence 21, Appl | C 671 | 13 | 2.2 | 1520 | 4 | US-09-273-565-20 | Sequence 20, Appl |
| 599 | 13 | 2.2 | 1134 | 1 | US-08-087-772A-14 | Sequence 14, Appl | C 672 | 13 | 2.2 | 1520 | 4 | US-09-100-664A-1 | Sequence 1, Appl |
| 600 | 13 | 2.2 | 1134 | 4 | US-09-206-059-29 | Sequence 29, Appl | 673 | 13 | 2.2 | 1524 | 1 | US-08-197-792-34 | Sequence 34, Appl |
| 601 | 13 | 2.2 | 1136 | 1 | US-08-106-981-5 | Sequence 5, Appl | 674 | 13 | 2.2 | 1524 | 1 | US-08-459-850-30 | Sequence 34, Appl |
| 602 | 13 | 2.2 | 1164 | 1 | US-08-087-772A-3 | Sequence 3, Appl | 675 | 13 | 2.2 | 1524 | 1 | US-08-459-850-30 | Sequence 34, Appl |
| 603 | 13 | 2.2 | 1185 | 1 | US-08-493-638-5 | Sequence 5, Appl | C 676 | 13 | 2.2 | 1524 | 6 | 512699-1 | Sequence 34, Appl |
| 604 | 13 | 2.2 | 1212 | 2 | US-08-802-322-4 | Sequence 4, Appl | C 677 | 13 | 2.2 | 1525 | 6 | 512699-1 | Sequence 34, Appl |
| 605 | 13 | 2.2 | 1215 | 2 | US-08-522-841-7 | Sequence 7, Appl | C 678 | 13 | 2.2 | 1529 | 3 | US-08-858-0876A-3 | Sequence 3, Appl |
| 606 | 13 | 2.2 | 1216 | 2 | US-08-522-841-3 | Sequence 3, Appl | 679 | 13 | 2.2 | 1537 | 2 | US-08-839-008-1 | Sequence 1, Appl |
| 607 | 13 | 2.2 | 1227 | 1 | US-07-916-901-1 | Sequence 1, Appl | 680 | 13 | 2.2 | 1560 | 1 | US-07-669-171-1 | Sequence 1, Appl |
| 608 | 13 | 2.2 | 1227 | 1 | US-08-351-473B-7 | Sequence 7, Appl | 681 | 13 | 2.2 | 1561 | 1 | US-07-824-247-44 | Sequence 3, Appl |
| 609 | 13 | 2.2 | 1236 | 1 | US-08-634-331-5 | Sequence 5, Appl | 682 | 13 | 2.2 | 1561 | 1 | US-08-142-473A-4 | Sequence 4, Appl |
| 610 | 13 | 2.2 | 1237 | 1 | US-08-240-372-1 | Sequence 1, Appl | 683 | 13 | 2.2 | 1561 | 1 | US-08-469-203A-4 | Sequence 4, Appl |
| 611 | 13 | 2.2 | 1240 | 1 | US-08-388-756-1 | Sequence 1, Appl | 684 | 13 | 2.2 | 1561 | 1 | US-08-469-203A-4 | Sequence 4, Appl |

| | | | | | | | | | | | | | |
|-------|----|-----|------|---|-------------------|--------------------|-------|----|-----|------|---|--------------------|--------------------|
| 685 | 13 | 2.2 | 1551 | 1 | US-08-469-203A-4 | Sequence 4, Appl1 | c 758 | 13 | 2.2 | 1886 | 3 | US-08-344-639E-5 | Sequence 5, Appl1 |
| 686 | 13 | 2.2 | 1561 | 3 | US-08-470-204A-44 | Sequence 44, Appl1 | c 759 | 13 | 2.2 | 1886 | 3 | US-09-262-927-1 | Sequence 1, Appl1 |
| c 687 | 13 | 2.2 | 1563 | 3 | US-08-738-168B-11 | Sequence 11, Appl1 | c 760 | 13 | 2.2 | 1886 | 4 | US-08-467-969A-5 | Sequence 5, Appl1 |
| 688 | 13 | 2.2 | 1569 | 1 | US-07-669-171-3 | Sequence 3, Appl1 | c 761 | 13 | 2.2 | 1886 | 4 | US-08-467-961A-5 | Sequence 5, Appl1 |
| 689 | 13 | 2.2 | 1576 | 1 | US-08-736-361A-1 | Sequence 1, Appl1 | c 762 | 13 | 2.2 | 1886 | 4 | US-08-001-554A-5 | Sequence 5, Appl1 |
| c 690 | 13 | 2.2 | 1581 | 2 | US-08-845-998-1 | Sequence 1, Appl1 | c 763 | 13 | 2.2 | 1889 | 3 | US-09-354-129-7 | Sequence 7, Appl1 |
| c 691 | 13 | 2.2 | 1581 | 2 | US-09-206-537-1 | Sequence 1, Appl1 | c 764 | 13 | 2.2 | 1902 | 3 | US-09-041-991A-7 | Sequence 9, Appl1 |
| c 692 | 13 | 2.2 | 1593 | 2 | US-08-524-828-2 | Sequence 2, Appl1 | c 765 | 13 | 2.2 | 1902 | 3 | US-09-041-991A-9 | Sequence 9, Appl1 |
| c 693 | 13 | 2.2 | 1593 | 3 | US-08-975-114A-2 | Sequence 2, Appl1 | c 766 | 13 | 2.2 | 1904 | 2 | US-08-559-939-3 | Sequence 3, Appl1 |
| c 694 | 13 | 2.2 | 1594 | 3 | US-08-933-750C-59 | Sequence 59, Appl1 | c 767 | 13 | 2.2 | 1904 | 2 | US-08-559-941-3 | Sequence 3, Appl1 |
| 695 | 13 | 2.2 | 1594 | 3 | US-08-933-750C-59 | Sequence 59, Appl1 | c 768 | 13 | 2.2 | 1904 | 3 | US-08-896-500-3 | Sequence 3, Appl1 |
| 696 | 13 | 2.2 | 1594 | 3 | US-09-234-613-59 | Sequence 50, Appl1 | c 769 | 13 | 2.2 | 1904 | 3 | US-08-476-397-3 | Sequence 3, Appl1 |
| 697 | 13 | 2.2 | 1600 | 2 | US-08-933-750C-50 | Sequence 50, Appl1 | c 770 | 13 | 2.2 | 1904 | 3 | US-08-973-720-3 | Sequence 3, Appl1 |
| 698 | 13 | 2.2 | 1600 | 3 | US-09-234-613-50 | Sequence 50, Appl1 | c 771 | 13 | 2.2 | 1904 | 3 | US-09-262-927-3 | Sequence 3, Appl1 |
| 699 | 13 | 2.2 | 1600 | 3 | US-08-524-828-3 | Sequence 3, Appl1 | c 772 | 13 | 2.2 | 1910 | 2 | US-09-009-438-1 | Sequence 1, Appl1 |
| 700 | 13 | 2.2 | 1608 | 2 | US-08-975-114A-3 | Sequence 3, Appl1 | c 773 | 13 | 2.2 | 1910 | 3 | US-09-207-493-1 | Sequence 1, Appl1 |
| 701 | 13 | 2.2 | 1608 | 3 | US-08-849-281A-1 | Sequence 1, Appl1 | c 774 | 13 | 2.2 | 1920 | 1 | US-08-687-772A-1 | Sequence 1, Appl1 |
| 702 | 13 | 2.2 | 1615 | 4 | US-08-308-814-1 | Sequence 1, Appl1 | c 775 | 13 | 2.2 | 1923 | 1 | US-08-077-939-14 | Sequence 14, Appl1 |
| c 703 | 13 | 2.2 | 1618 | 1 | US-08-410-540-1 | Sequence 1, Appl1 | c 776 | 13 | 2.2 | 1923 | 1 | US-08-461-599-14 | Sequence 14, Appl1 |
| 704 | 13 | 2.2 | 1629 | 1 | US-08-375-709-16 | Sequence 16, Appl1 | c 777 | 13 | 2.2 | 1923 | 1 | US-08-461-621-14 | Sequence 14, Appl1 |
| 705 | 13 | 2.2 | 1629 | 1 | US-08-752-929-16 | Sequence 16, Appl1 | c 778 | 13 | 2.2 | 1923 | 1 | US-08-465-334-14 | Sequence 14, Appl1 |
| 706 | 13 | 2.2 | 1635 | 1 | US-08-499-568-14 | Sequence 14, Appl1 | c 779 | 13 | 2.2 | 1926 | 2 | US-08-464-517-5 | Sequence 5, Appl1 |
| 707 | 13 | 2.2 | 1635 | 1 | US-08-793-958-14 | Sequence 14, Appl1 | c 780 | 13 | 2.2 | 1926 | 5 | US-08-463-772-5 | Sequence 5, Appl1 |
| 708 | 13 | 2.2 | 1635 | 2 | US-08-956-998-1 | Sequence 1, Appl1 | c 781 | 13 | 2.2 | 1926 | 5 | PCT-US93-05000-5 | Sequence 9, Appl1 |
| c 709 | 13 | 2.2 | 1641 | 4 | US-08-659-254-19 | Sequence 19, Appl1 | c 782 | 13 | 2.2 | 1928 | 3 | US-09-008-481A-9 | Sequence 9, Appl1 |
| c 710 | 13 | 2.2 | 1644 | 3 | US-09-339-993-3 | Sequence 3, Appl1 | c 783 | 13 | 2.2 | 1928 | 3 | US-09-195-666A-15 | Sequence 15, Appl1 |
| 711 | 13 | 2.2 | 1666 | 2 | US-08-666-367B-3 | Sequence 3, Appl1 | c 784 | 13 | 2.2 | 1930 | 1 | US-07-982-112-1 | Sequence 1, Appl1 |
| 712 | 13 | 2.2 | 1666 | 4 | US-09-143-438-3 | Sequence 3, Appl1 | c 785 | 13 | 2.2 | 1941 | 2 | US-09-008-960-2 | Sequence 2, Appl1 |
| 713 | 13 | 2.2 | 1680 | 3 | US-08-891-845-3 | Sequence 3, Appl1 | c 786 | 13 | 2.2 | 1941 | 3 | US-09-368-240-2 | Sequence 2, Appl1 |
| 714 | 13 | 2.2 | 1680 | 3 | US-08-869-696-8 | Sequence 8, Appl1 | c 787 | 13 | 2.2 | 1947 | 2 | US-08-989-825-2 | Sequence 2, Appl1 |
| c 715 | 13 | 2.2 | 1690 | 3 | US-08-869-696-8 | Sequence 8, Appl1 | c 788 | 13 | 2.2 | 1960 | 2 | US-08-415-823-3 | Sequence 3, Appl1 |
| 716 | 13 | 2.2 | 1695 | 3 | US-09-008-481A-3 | Sequence 3, Appl1 | c 789 | 13 | 2.2 | 1960 | 2 | US-09-086-662-3 | Sequence 3, Appl1 |
| 717 | 13 | 2.2 | 1695 | 3 | US-09-195-666A-17 | Sequence 17, Appl1 | c 790 | 13 | 2.2 | 1972 | 3 | US-08-564-110C-1 | Sequence 1, Appl1 |
| c 718 | 13 | 2.2 | 1719 | 3 | US-09-160-065-1 | Sequence 1, Appl1 | c 791 | 13 | 2.2 | 1974 | 1 | US-08-077-939-16 | Sequence 16, Appl1 |
| c 719 | 13 | 2.2 | 1722 | 1 | US-08-261-822A-9 | Sequence 9, Appl1 | c 792 | 13 | 2.2 | 1974 | 1 | US-08-461-599-16 | Sequence 16, Appl1 |
| c 720 | 13 | 2.2 | 1722 | 2 | US-08-467-963C-27 | Sequence 27, Appl1 | c 793 | 13 | 2.2 | 1974 | 1 | US-08-461-621-16 | Sequence 16, Appl1 |
| c 721 | 13 | 2.2 | 1722 | 2 | US-08-838-189D-27 | Sequence 27, Appl1 | c 794 | 13 | 2.2 | 1974 | 1 | US-08-465-334-16 | Sequence 16, Appl1 |
| c 722 | 13 | 2.2 | 1722 | 3 | US-08-852-344D-27 | Sequence 27, Appl1 | c 795 | 13 | 2.2 | 1992 | 4 | US-09-276-531-19 | Sequence 4, Appl1 |
| c 723 | 13 | 2.2 | 1722 | 3 | US-08-344-639E-27 | Sequence 27, Appl1 | c 796 | 13 | 2.2 | 1996 | 1 | US-08-385-570-1 | Sequence 1, Appl1 |
| c 724 | 13 | 2.2 | 1722 | 3 | PCT-US95-07744A-9 | Sequence 9, Appl1 | c 797 | 13 | 2.2 | 2003 | 1 | US-08-036-555E-21 | Sequence 21, Appl1 |
| c 725 | 13 | 2.2 | 1729 | 2 | US-08-844-024-1 | Sequence 1, Appl1 | c 798 | 13 | 2.2 | 2003 | 1 | US-08-469-569-21 | Sequence 21, Appl1 |
| c 726 | 13 | 2.2 | 1730 | 3 | US-08-718-547-1 | Sequence 1, Appl1 | c 799 | 13 | 2.2 | 2003 | 1 | US-08-216-219-1 | Sequence 1, Appl1 |
| c 727 | 13 | 2.2 | 1730 | 3 | US-09-222-817-11 | Sequence 13, Appl1 | c 800 | 13 | 2.2 | 2003 | 1 | US-08-249-322A-21 | Sequence 21, Appl1 |
| c 728 | 13 | 2.2 | 1730 | 3 | US-09-222-817-13 | Sequence 13, Appl1 | c 801 | 13 | 2.2 | 2003 | 1 | US-08-469-526A-21 | Sequence 21, Appl1 |
| c 729 | 13 | 2.2 | 1762 | 6 | 5185262-2 | Patent No. 5185262 | c 802 | 13 | 2.2 | 2003 | 2 | US-08-734-591A-21 | Sequence 21, Appl1 |
| 730 | 13 | 2.2 | 1763 | 6 | 5185262-1 | Sequence 1, Appl1 | c 803 | 13 | 2.2 | 2003 | 3 | US-08-469-660-21 | Sequence 21, Appl1 |
| c 731 | 13 | 2.2 | 1766 | 5 | PCT-US93-00601-1 | Sequence 1, Appl1 | c 804 | 13 | 2.2 | 2003 | 3 | US-08-341-018-71 | Sequence 71, Appl1 |
| c 732 | 13 | 2.2 | 1768 | 5 | PCT-US94-07107A-1 | Sequence 5, Appl1 | c 805 | 13 | 2.2 | 2003 | 3 | US-08-470-339-21 | Sequence 21, Appl1 |
| 733 | 13 | 2.2 | 1780 | 3 | US-08-933-821-5 | Sequence 5, Appl1 | c 806 | 13 | 2.2 | 2003 | 4 | US-08-735-021-21 | Sequence 21, Appl1 |
| 734 | 13 | 2.2 | 1780 | 3 | US-08-960-507-5 | Sequence 5, Appl1 | c 807 | 13 | 2.2 | 2003 | 4 | US-08-734-664A-21 | Sequence 21, Appl1 |
| 735 | 13 | 2.2 | 1809 | 1 | US-08-455-001-1 | Sequence 1, Appl1 | c 808 | 13 | 2.2 | 2003 | 4 | US-08-470-339-21 | Sequence 21, Appl1 |
| 736 | 13 | 2.2 | 1809 | 5 | PCT-US95-11865-1 | Sequence 1, Appl1 | c 809 | 13 | 2.2 | 2003 | 5 | PCT-US93-00537A-1 | Sequence 21, Appl1 |
| c 737 | 13 | 2.2 | 1815 | 3 | US-09-041-545-1 | Sequence 1, Appl1 | c 810 | 13 | 2.2 | 2003 | 5 | PCT-US94-05083C-21 | Sequence 21, Appl1 |
| c 738 | 13 | 2.2 | 1815 | 3 | US-09-327-925-1 | Sequence 1, Appl1 | c 811 | 13 | 2.2 | 2003 | 5 | PCT-US94-05083C-21 | Sequence 21, Appl1 |
| 739 | 13 | 2.2 | 1826 | 2 | US-08-808-931-15 | Sequence 15, Appl1 | c 812 | 13 | 2.2 | 2009 | 1 | US-08-333-358-9 | Sequence 9, Appl1 |
| 740 | 13 | 2.2 | 1826 | 3 | US-08-808-931-15 | Sequence 15, Appl1 | c 813 | 13 | 2.2 | 2009 | 1 | US-08-463-694-9 | Sequence 9, Appl1 |
| 741 | 13 | 2.2 | 1826 | 3 | US-09-050-603A-15 | Sequence 15, Appl1 | c 814 | 13 | 2.2 | 2009 | 1 | US-08-694-501-9 | Sequence 9, Appl1 |
| 742 | 13 | 2.2 | 1826 | 3 | US-09-102-420B-15 | Sequence 15, Appl1 | c 815 | 13 | 2.2 | 2009 | 2 | US-08-484-200-3 | Sequence 3, Appl1 |
| 743 | 13 | 2.2 | 1830 | 2 | US-09-102-420B-15 | Sequence 15, Appl1 | c 816 | 13 | 2.2 | 2022 | 2 | US-08-484-200-3 | Sequence 3, Appl1 |
| 744 | 13 | 2.2 | 1830 | 3 | US-08-933-750C-79 | Sequence 79, Appl1 | c 817 | 13 | 2.2 | 2022 | 5 | PCT-US96-00996-4 | Sequence 4, Appl1 |
| c 745 | 13 | 2.2 | 1884 | 1 | US-09-234-613-79 | Sequence 79, Appl1 | c 818 | 13 | 2.2 | 2027 | 1 | US-08-150-203A-1 | Sequence 1, Appl1 |
| c 746 | 13 | 2.2 | 1884 | 1 | US-08-307-382-3 | Sequence 3, Appl1 | c 819 | 13 | 2.2 | 2027 | 1 | US-08-454-730-1 | Sequence 1, Appl1 |
| c 747 | 13 | 2.2 | 1884 | 1 | US-08-366-779-3 | Sequence 3, Appl1 | c 820 | 13 | 2.2 | 2027 | 2 | US-08-377-309-1 | Sequence 1, Appl1 |
| c 748 | 13 | 2.2 | 1884 | 1 | US-08-478-727-3 | Sequence 3, Appl1 | c 821 | 13 | 2.2 | 2033 | 1 | US-08-377-309-1 | Sequence 1, Appl1 |
| c 749 | 13 | 2.2 | 1884 | 1 | US-08-473-508-3 | Sequence 3, Appl1 | c 822 | 13 | 2.2 | 2033 | 1 | US-08-148-910-14 | Sequence 14, Appl1 |
| c 750 | 13 | 2.2 | 1886 | 2 | US-08-789-936-3 | Sequence 3, Appl1 | c 823 | 13 | 2.2 | 2033 | 1 | US-08-448-937A-14 | Sequence 14, Appl1 |
| c 751 | 13 | 2.2 | 1886 | 2 | US-08-659-939-1 | Sequence 1, Appl1 | c 824 | 13 | 2.2 | 2033 | 1 | US-08-448-937A-14 | Sequence 14, Appl1 |
| c 752 | 13 | 2.2 | 1886 | 2 | US-08-850-041-1 | Sequence 1, Appl1 | c 825 | 13 | 2.2 | 2034 | 3 | US-08-834-314-1 | Sequence 14, Appl1 |
| c 753 | 13 | 2.2 | 1886 | 2 | US-08-467-963C-5 | Sequence 5, Appl1 | c 826 | 13 | 2.2 | 2043 | 3 | US-07-737-736B-6 | Sequence 6, Appl1 |
| c 754 | 13 | 2.2 | 1886 | 2 | US-08-838-189D-5 | Sequence 5, Appl1 | c 827 | 13 | 2.2 | 2054 | 1 | US-08-351-981-1 | Sequence 1, Appl1 |
| c 755 | 13 | 2.2 | 1886 | 2 | US-08-952-344D-5 | Sequence 5, Appl1 | c 828 | 13 | 2.2 | 2057 | 1 | US-08-450-945-57 | Sequence 57, Appl1 |
| c 756 | 13 | 2.2 | 1886 | 3 | US-08-896-500-1 | Sequence 1, Appl1 | c 829 | 13 | 2.2 | 2057 | 4 | US-08-976-161-57 | Sequence 57, Appl1 |
| c 757 | 13 | 2.2 | 1886 | 3 | US-08-476-397-1 | Sequence 1, Appl1 | c 830 | 13 | 2.2 | 2070 | 1 | US-08-233-146-7 | Sequence 7, Appl1 |

| | | | | | | | |
|--------|----|-----|------|---|------------------|--------------|-------|
| C 977 | 13 | 2.2 | 2729 | 5 | PCr-US93-08131-1 | Sequence 1, | Appl1 |
| C 978 | 13 | 2.2 | 2732 | 1 | US-08-476-000-60 | Sequence 60, | Appl1 |
| C 979 | 13 | 2.2 | 2732 | 1 | US-08-472-840-60 | Sequence 60, | Appl1 |
| C 980 | 13 | 2.2 | 2732 | 2 | US-08-476-976-60 | Sequence 60, | Appl1 |
| C 981 | 13 | 2.2 | 2732 | 2 | US-08-476-976-60 | Sequence 60, | Appl1 |
| C 982 | 13 | 2.2 | 2745 | 5 | US-08-385-370-3 | Sequence 3, | Appl1 |
| C 983 | 13 | 2.2 | 2745 | 5 | PCr-US94-03705-3 | Sequence 3, | Appl1 |
| C 984 | 13 | 2.2 | 2746 | 2 | US-08-576-165-3 | Sequence 3, | Appl1 |
| C 985 | 13 | 2.2 | 2752 | 1 | US-08-430-925A-3 | Sequence 3, | Appl1 |
| C 986 | 13 | 2.2 | 2753 | 1 | US-07-854-603-1 | Sequence 1, | Appl1 |
| C 987 | 13 | 2.2 | 2770 | 3 | US-08-586-165-6 | Sequence 1, | Appl1 |
| C 988 | 13 | 2.2 | 2774 | 4 | US-09-363-189B-5 | Sequence 6, | Appl1 |
| C 989 | 13 | 2.2 | 2793 | 2 | US-08-795-866-13 | Sequence 5, | Appl1 |
| C 990 | 13 | 2.2 | 2796 | 1 | US-08-261-677-8 | Sequence 13, | Appl1 |
| C 991 | 13 | 2.2 | 2796 | 1 | US-08-384-556A-4 | Sequence 8, | Appl1 |
| C 992 | 13 | 2.2 | 2796 | 2 | US-08-331-355A-8 | Sequence 4, | Appl1 |
| C 993 | 13 | 2.2 | 2796 | 3 | US-08-777-147-5 | Sequence 8, | Appl1 |
| C 994 | 13 | 2.2 | 2796 | 3 | US-09-157-077-8 | Sequence 5, | Appl1 |
| C 995 | 13 | 2.2 | 2796 | 5 | PCr-US94-12334-8 | Sequence 8, | Appl1 |
| C 996 | 13 | 2.2 | 2796 | 5 | PCr-US95-07753-4 | Sequence 8, | Appl1 |
| C 997 | 13 | 2.2 | 2799 | 3 | US-08-968-752B-5 | Sequence 4, | Appl1 |
| C 998 | 13 | 2.2 | 2808 | 5 | US-08-237-919-1 | Sequence 1, | Appl1 |
| C 999 | 13 | 2.2 | 2808 | 5 | PCr-US95-05518-1 | Sequence 1, | Appl1 |
| C 1000 | 13 | 2.2 | 2830 | 1 | US-07-882-292-1 | Sequence 1, | Appl1 |

ALIGNMENTS

```

RESULT 1
US-08-480-917-1
: Sequence 1, Application US/08480917
: Patent No. 5820864
: GENERAL INFORMATION:
: APPLICANT: PARANHOS-BACCALA, Glaucia
: APPLICANT: LESENECHAL, Mylene
: APPLICANT: JOLIVET, Michel
: TITLE OF INVENTION: NEW TRYPAOSOMA CRUZI ANTIGEN, AND GENE
: TITLE OF INVENTION: ENCODING THE LATTER, THEIR APPLICATION TO THE DETECTION OF
: TITLE OF INVENTION: CHAGAS DISEASE
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Oliff & Berridge
: STREET: 700 South Washington Street, Suite 300
: CITY: Alexandria
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22314
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/480,917
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Berridge, William P.
: REGISTRATION NUMBER: 30,024
: REFERENCE/DOCKET NUMBER: WPB 36400
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-836-6400
: TELEFAX: 703-836-2787
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3402 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: US-08-480-917-1

```

| | | | | |
|-----------------------|-----------------|---------------------|-----------|--------------|
| Query Match | 100.0%; | Score 594; | DB 1; | Length 3402; |
| Best Local Similarity | 100.0%; | Pred. No. 4.5e+297; | | |
| Matches 594; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0 |

| | | | |
|----|------|---|------|
| Oy | 1 | caagttacagcgttaacagccttttcttcaatcgttaacgcgcgcgttagctcgcttccgt | 60 |
| Db | 1232 | CAGGTACAGCCTTAACGGCTTTTGCTTTCATCTGACAGCCGACGGTAAGCTGCTCTGGCT | 1291 |
| Oy | 61 | gacatgtcgaattcgaatcgagcatctccatctccgttccctccgcagggaaagaaacagag | 120 |
| Db | 1292 | GACATGTGCAATTCGATTGAGGATCTTCGATCTCCGGTCCCTCCGACGGGAAGAACAGCAG | 1351 |
| Oy | 121 | ccagggccaaaaaacaatcgtgtagtctgcgaacgcgcgaacccgggtgtgtgtctccgcgac | 180 |
| Db | 1352 | CCAGGCCAAAAAATCGTGTGATGTGGCGACGGCAAAACCGGGGTGTGTGTCTCGGGGCACT | 1411 |
| Oy | 181 | gagccgcgcgagttagcagtaaccataagacctctgcgcgtgcgtgcgaatccctctacata | 240 |
| Db | 1412 | GAGCCGGCGAGTACACTCATACCAATACACTTCTGGCGCTGCTGCATCCCTCATCA | 1471 |
| Oy | 241 | cccccttcttaagcgcagcacaagcagcgcgcctctctgcgcgcgcgcgtatcgcgtctag | 300 |
| Db | 1472 | CCCCCTGTTTCAGCGGCCAAGCAGCAGCCGGCCTCTGCCCCGGCGGATCGGGCTAG | 1531 |
| Oy | 301 | ccgcacgttggggagcaagatcattgtctaactctagtgaatcagctgggataatgatac | 360 |
| Db | 1532 | CCGCACGTGGGGAGACCAAGATCATTTGCTAACTGTAGTAATACACTGGGCAATTAAATGCAC | 1591 |
| Oy | 361 | caagagagcgtctgctcaacacattgtagagccgcgcacagaaggtcttaagcgcgttagcgtcc | 420 |
| Db | 1592 | CAAAAGGAGCGTGTACACTGTGAAGCCCGCCGCACACAGAGCTTACGGCGGTGACGTCC | 1651 |
| Oy | 421 | acgaatcacgcgccgcgcagcgaacaagtctcaatcagggacaaatctgcgcgaactctgaacgt | 480 |
| Db | 1652 | ACGACTACCGCCCGCAGCGCAACATGTCCTACGGGACAAATGGCCGAGACTTGAAGGCT | 1711 |
| Oy | 481 | ggattgtctgcagctaaatagtgtgcgcgcgcgcgcctcgtctctccaaagcgcgcgcgaa | 540 |
| Db | 1712 | GGATTGTGCGACTAATTAATGTGGGCGAGCGCGGCTCTGCTCCACACCCCGCGGAA | 1771 |
| Oy | 541 | ccaacaggaagaanaagcctccgcgcgcatctggaacagctccgtgtgcata | 594 |
| Db | 1772 | CCAACAGGAGAAAGAAAGGCTCTCCGCGCATGTGAAACAGAGCTCCCTGTGGCGATA | 1825 |

RESULT 2
 US-08-480-917-9
 ; Sequence 9, Application US/08480917
 ; Patent No. 5820864
 ; GENERAL INFORMATION:
 ; APPLICANT: PARANHOS-BACCALA, Glaucia
 ; APPLICANT: LESENECHAL, Mylene
 ; APPLICANT: JOLIVET, Michel
 ; TITLE OF INVENTION: NEW TRYPANOSOMA CRUZI ANTIGEN, AND GENE
 ; TITLE OF INVENTION: ENCODING THE LATTER, THEIR APPLICATION TO THE DETECTION OF
 ; TITLE OF INVENTION: CHAGAS DISEASE
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Oliff & Bertridge
 ; STREET: 700 South Washington Street, Suite 300
 ; CITY: Alexandria
 ; STATE: Virginia
 ; COUNTRY: U.S.A.
 ; ZIP: 22314
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/480,917
 ; FILING DATE: 07-JUN-1995

```
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-480-917-9

Query Match      3.7%; Score 22; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 cagccgacgtagctgcgtcct 56
Db 1 CAGCCGACGCTAGCTGCTGCT 22

RESULT 3
US-08-970-269A-32/c
; Sequence 32, Application US/08970269A
; Patent No. 5976803
; GENERAL INFORMATION:
; APPLICANT: Kathryn Meek
; TITLE OF INVENTION: Genetic Test For Equine Severe
; TITLE OF INVENTION: Combined Immunodeficiency Disease
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Benjamin A. Adler
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/970/269A
; FILING DATE: No. 5976803ember 14, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler Ph.D., Benjamin A.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5860
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-777-2321
; TELEFAX: 713-777-6908
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11873 bp
; TYPE: nucleic acid
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: other nucleic acid
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; FEATURE:
; US-08-970-269A-32
```

```
Query Match      3.2%; Score 19; DB 2; Length 11873;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 211 actctgcgcgtcgtcat 229
Db 6446 ACTCTGCCGCTGCTCAT 6428

RESULT 4
US-08-970-269A-31/c
; Sequence 31, Application US/08970269A
; Patent No. 5976803
; GENERAL INFORMATION:
; APPLICANT: Kathryn Meek
; TITLE OF INVENTION: Genetic Test For Equine Severe
; TITLE OF INVENTION: Combined Immunodeficiency Disease
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Benjamin A. Adler
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/970/269A
; FILING DATE: No. 5976803ember 14, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler Ph.D., Benjamin A.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5860
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-777-2321
; TELEFAX: 713-777-6908
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11878 bp
; TYPE: nucleic acid
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: other nucleic acid
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; FEATURE:
; US-08-970-269A-31

Query Match      3.2%; Score 19; DB 2; Length 11878;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 211 actctgcgcgtcgtcat 229
Db 6446 ACTCTGCCGCTGCTCAT 6428

RESULT 5
US-08-970-269A-28/c
; Sequence 28, Application US/08970269A
; Patent No. 5976803
; GENERAL INFORMATION:
; APPLICANT: Kathryn Meek
; TITLE OF INVENTION: Genetic Test For Equine Severe
```

```

; TITLE OF INVENTION: Combined Immunodeficiency Disease
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Benjamin A. Adler
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/970,269A
; FILING DATE: No. 5976803ember 14, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler Ph.D., Benjamin A.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5860
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-777-2321
; TELEFAX: 713-777-6908
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11883 bp
; TYPE: nucleic acid
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: other nucleic acid
; HYPOTHEICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; US-08-970-269A-28

Query Match      3.2%; Score 19; DB 2; Length 11883;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 actctgcgcgtctgcac 229
|||||
Db 6449 ACTCTGCCGCTGCTGCAT 6431

RESULT 6
US-08-480-917-5
; Sequence 5, Application US/08480917
; Patent No. 5820864
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: LESENECHAL, Mylene
; APPLICANT: JOLIVET, Michel
; TITLE OF INVENTION: NEW TRYPAPOSOMA CRUZI ANTIGEN, AND GENE
; TITLE OF INVENTION: ENCODING THE LATTER; THEIR APPLICATION TO THE DETECTION OF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/480,917
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: MPB 36400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-480-917-5

```

```

Query Match      3.0%; Score 18; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 tcgggcactgacgcgcgc 189
|||||
Db 1 TCGGCGACTGACGCGCGC 18

```

```

RESULT 7
US-08-480-917-8/c
; Sequence 8, Application US/08480917
; Patent No. 5820864
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: LESENECHAL, Mylene
; APPLICANT: JOLIVET, Michel
; TITLE OF INVENTION: NEW TRYPAPOSOMA CRUZI ANTIGEN, AND GENE
; TITLE OF INVENTION: ENCODING THE LATTER; THEIR APPLICATION TO THE DETECTION OF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,917
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: MPB 36400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-480-917-8

```

Query Match 3.0%; Score 18; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 actctgcgcctcctcga 228
|||||
DB 18 ACTCTGCCGCTGCTGCA 1

RESULT 8
US-08-454-557C-50
; Sequence 50, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609,3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2520 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both

US-08-454-557C-50

Query Match 3.0%; Score 18; DB 2; Length 2520;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 cccctgcatacccccctg 247
|||||
DB 298 CCCCTGCATCACCCTCG 315

RESULT 9
US-08-340-426D-50
; Sequence 50, Application US/08340426D
; Patent No. 5948634
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609,3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2520 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both

US-08-340-426D-50

Query Match 3.0%; Score 18; DB 2; Length 2520;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 cccctgcatacccccctg 247
|||||
DB 298 CCCCTGCATCACCCTCG 315

RESULT 10
US-08-450-673C-50
; Sequence 50, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609,3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 50:

SEQUENCE CHARACTERISTICS:
LENGTH: 2520 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-50

Query Match 3.0%; Score 18; DB 2; Length 2520;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 cccctgcataccccctg 247
|||||
DB 298 CCCCTGCATCACCCTG 315

RESULT 11
PCT-US95-17111A-50
Sequence 50, Application PC/TUS9517111A

GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and
TITLE OF INVENTION: Detection of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 2520 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-50

Query Match 3.0%; Score 18; DB 5; Length 2520;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 cccctgcataccccctg 247
|||||
DB 298 CCCCTGCATCACCCTG 315

RESULT 12
US-09-335-409-1/C
Sequence 1, Application US/09335409

Patent No. 6121029
GENERAL INFORMATION:
APPLICANT: Schnupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/335,409
CURRENT FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 68750
TYPE: DNA
ORGANISM: Sorangium cellulosum
US-09-335-409-1

Query Match 3.0%; Score 18; DB 3; Length 68750;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 517 tgcgtccacagccgcg 534
|||||
DB 19020 TCGTCTCCACAGCCGCG 19003

RESULT 13
US-08-692-787-4
Sequence 4, Application US/08692787
Patent No. 5882864
GENERAL INFORMATION:
APPLICANT: An, Gang
APPLICANT: O'Hara, S. Mark
APPLICANT: Ralph, David
APPLICANT: Velltri, Robert
TITLE OF INVENTION: BIOMARKERS AND TARGETS FOR DIAGNOSIS,
TITLE OF INVENTION: PROGNOSIS AND MANAGEMENT OF PROSTATE
DISEASE
NUMBER OF SEQUENCES: 82
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/692,787
FILING DATE: Concurrently Herewith
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Cordier, Timothy S.
REGISTRATION NUMBER: 38,414
REFERENCE/DOCKET NUMBER: UROC:012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 673 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-692-787-4

Query Match 2.9%; Score 17; DB 2; Length 673;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaaccacagagagaag 553
|||||
Db 542 GAACCAACAGAGAGAG 558

RESULT 14
US-09-097-199-4
; Sequence 4, Application US/09097199
; Patent No. 6218529
; GENERAL INFORMATION:
; APPLICANT: An, Gang
; APPLICANT: O'Hara, S. Mark
; APPLICANT: Ralph, David
; APPLICANT: Veltri, Robert
; TITLE OF INVENTION: BIOMARKERS AND TARGETS FOR DIAGNOSIS,
; TITLE OF INVENTION: PROGNOSIS AND MANAGEMENT OF PROSTATE DISEASE
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,199
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/692,787
; FILING DATE: 31-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nakashima, Richard A.
; REGISTRATION NUMBER: P-42,023
; REFERENCE/DOCKET NUMBER: UROC:018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 673 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-097-199-4

Query Match 2.9%; Score 17; DB 4; Length 673;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaaccacagagagaag 553
|||||
Db 542 GAACCAACAGAGAGAG 558

RESULT 15
US-08-072-610-1
; Sequence 1, Application US/08072610
; Patent No. 5532133
; GENERAL INFORMATION:
; APPLICANT: Barnwell, John

; TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens,
; TITLE OF INVENTION: Monoclonal Antibodies, and Diagnostic Assays
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby and Darby
; STREET: 805 Third Ave.
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022-7513

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,610
; FILING DATE: 19930602
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gogoris, Adda
; REGISTRATION NUMBER: 29,714
; REFERENCE/DOCKET NUMBER: 5986/07686
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)527-7700
; TELEFAX: (212)753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3337 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium vivax
; IMMEDIATE SOURCE:
; CLONE: PvMB3.3.1
; US-08-072-610-1

Query Match 2.9%; Score 17; DB 1; Length 3337;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 aaccacagagagaaga 555
|||||
Db 2029 AACCAACAGAGAGAGA 2045

RESULT 16
US-08-719-822B-1
; Sequence 1, Application US/08719822B
; Patent No. 5874527
; GENERAL INFORMATION:
; APPLICANT: Barnwell, John
; TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby and Darby
; STREET: 805 Third Ave.
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022-7513
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/719,822B
FILING DATE: 09/30/96
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gogoris, Adda
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 5986/17686US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)527-7700
TELEFAX: (212)753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3337 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium vivax
IMMEDIATE SOURCE:
CLONE: PVM3.3.1
US-08-719-822B-1

Query Match 2.9%; Score 17; DB 2; Length 3337;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 539 aaccaacagagaaga 555
|||||
Db 2029 AACCAACAGAGAAGAA 2045

RESULT 17
US-09-092-458-1
Sequence 1, Application US/09092458
Patent No. 6231861
GENERAL INFORMATION:
APPLICANT: Barnwell, John
TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens,
TITLE OF INVENTION: Monoclonal Antibodies, and Diagnostic Assays
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby and Darby
STREET: 805 Third Ave.
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022-7513
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/092,458
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/719,821
FILING DATE: 09/30/96
ATTORNEY/AGENT INFORMATION:
NAME: Gogoris, Adda
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 5986/17686US3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)527-7700
TELEFAX: (212)753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 3337 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium vivax
IMMEDIATE SOURCE:
CLONE: PVM3.3.1
US-09-092-458-1

Query Match 2.9%; Score 17; DB 4; Length 3337;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 539 aaccaacagagaaga 555
|||||
Db 2029 AACCAACAGAGAAGAA 2045

RESULT 18
US-08-142-897-1/C
Sequence 1, Application US/08142897
Patent No. 5447852
GENERAL INFORMATION:
APPLICANT: Friedman, Jeffrey S.
APPLICANT: Weissman, Irving L.
TITLE OF INVENTION: No. 5447852el Cyclophilins, Associating Proteins
TITLE OF INVENTION: and Uses
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Tracy J. Dunn
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142,897
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/005,917
FILING DATE: 15-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/740,375
FILING DATE: 05-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Dunn, Tracy D.
REGISTRATION NUMBER: 34,587
REFERENCE/DOCKET NUMBER: 5490A-92-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Primer
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..41

OTHER INFORMATION: /note="Adaptor-Primer"
US-08-142-897-1

Query Match 2.7%; Score 16; DB 1; Length 41;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 554 aaagcgctcgcgcc 569
Db 27 AAAGGCTCGCGCC 12

RESULT 19

US-08-628-039-9/c
Sequence 9, Application US/08628039
Patent No. 5942660
GENERAL INFORMATION:
APPLICANT: Gruys, Kenneth J.
APPLICANT: Misky, Timothy A.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Slater, Steven C.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stark, David M.
APPLICANT: Hinchee, Maud A. W.
APPLICANT: Clemente, Thomas E.
APPLICANT: Connor-Ward, Dannelte V.
APPLICANT: Fedele, Mary J.
APPLICANT: Fry, Joyce E.
APPLICANT: Howe, Arlene R.
APPLICANT: Rozman, Renee J.
TITLE OF INVENTION: Methods of Optimizing Substrate Pools and
TITLE OF INVENTION: Biosynthesis of poly-B-hydroxybutyrate-co-poly-B-hydroxyvalerate
TITLE OF INVENTION: In Bacteria and Plants
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gary M. Bond, Monsanto Company, A3SB
STREET: 800 No. 594260th Lindbergh Boulevard
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63167
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/628,039
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE: 13-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bond, Gary
REGISTRATION NUMBER: 29,283
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)694-3412
TELEFAX: (314)695-5435
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1185 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-628-039-9

Query Match 2.7%; Score 16; DB 2; Length 1185;
Best Local Similarity 100.0%; Pred. No. 39;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 506 ccagcgcgctcgctc 521
Db 544 CCAGCGCGCTCGCTC 529

RESULT 20

US-08-912-205-9/c
Sequence 9, Application US/08912205
Patent No. 6091002
GENERAL INFORMATION:
APPLICANT: Asrar, Jawed
APPLICANT: Misky, Timothy A.
APPLICANT: Shah, Devang T.
TITLE OF INVENTION: Polyhydroxyalkanoates of narrow molecular
TITLE OF INVENTION: Weight distribution prepared in transgenic plants
NUMBER OF SEQUENCES: 11
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/912,205
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/912,205
FILING DATE:
INFORMATION FOR SEQ ID NO: 9:

US-08-912-205-9

Query Match 2.7%; Score 16; DB 3; Length 1185;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 506 ccagcgcgctcgctc 521
Db 544 CCAGCGCGCTCGCTC 529

RESULT 21

US-09-440-400-9/c
Sequence 9, Application US/09440400
Patent No. 6228623
GENERAL INFORMATION:
APPLICANT: Asrar, Jawed
APPLICANT: Misky, Timothy A.
APPLICANT: Shah, Devang T.
TITLE OF INVENTION: Polyhydroxyalkanoates of narrow molecular
TITLE OF INVENTION: Weight distribution prepared in transgenic plants
NUMBER OF SEQUENCES: 11
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/440,400
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/912,205
FILING DATE:
INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:
LENGTH: 1185 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-440-400-9

Query Match 2.7%; Score 16; DB 4; Length 1185;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 506 ccagcgcgacctgtc 521
|||||
DB 544 CCAGCGCGCCTGCTC 529

RESULT 22

US-09-031-059-2
Sequence 2, Application US/09031059
Patent No. 5948659

GENERAL INFORMATION:

APPLICANT: KATO, NOBUO
APPLICANT: SAKAI, YASUYOSHI
APPLICANT: TANI, YOSHIKI
APPLICANT: FUKUYA, HIROSHI
TITLE OF INVENTION: RECOMBINANT FRUCTOSYL AMINO ACID
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSER: BIRCH, STEWART, KOLASCH AND BIRCH
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,059
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/899,172

ATTORNEY/AGENT INFORMATION:
FILING DATE:
NAME: WEINER, MARC S
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 0020-4253P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1314 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1311
US-09-031-059-2

Query Match 2.7%; Score 16; DB 2; Length 1314;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 351 taatgtcaccacaag 366

|||||
DB 1281 TAAATGTCACCAAGC 1296

RESULT 23

US-07-952-817-10
Sequence 10, Application US/07952817
Patent No. 5356796

GENERAL INFORMATION:

APPLICANT: KELLER, JOHN W.
TITLE OF INVENTION: A Repressor Protein and Gene for Regulating
TITLE OF INVENTION: Expression of Polypeptides and Its Use in the Preparation o
TITLE OF INVENTION: 2,2-Dialkylglycine Decarboxylase of Pseudomonas Cepacia
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSER: Finnegan, Henderson, Farabow, Garrett &
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: US
ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/952,817
FILING DATE: 19920928
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 01120.0002-01000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1411 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-952-817-10

Query Match 2.7%; Score 16; DB 1; Length 1411;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 459 caatggcgacctgtg 474
|||||
DB 372 CAATGGCGGACTGTG 387

RESULT 24

US-07-952-817-10
Sequence 10, Application US/07952817
Patent No. 5356796

GENERAL INFORMATION:

APPLICANT: KELLER, JOHN W.
TITLE OF INVENTION: REPRESSOR PROTEIN GENE FOR REGULATING
EXPRESSION OF POLYPEPTIDES AND ITS USE IN THE PREPARATION OF
2,2-DIALKYLGLYCINE DECARBOXYLASE OF PSEUDOMONAS CEPACIA
NUMBER OF SEQUENCES: 18
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/501,814
FILING DATE: 30-MAR-1990
SEQ ID NO: 3:
LENGTH: 1411
US-07-952-817-10

Query Match 2.7%; Score 16; DB 2; Length 1411;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5210025-3

Query Match 2.7%; Score 16; DB 6; Length 1411;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 459 caatgcccgcactgtg 474
|||||
Db 372 caatgcccgcactgtg 387

RESULT 25

US-08-360-6068-29/c
Sequence 29, Application US/08360606B
Patent No. 5919617
GENERAL INFORMATION:
APPLICANT: Jnanendra K. Bhattacharjee
APPLICANT: Richard C. Garrad
APPLICANT: Paul L. Skatrud
APPLICANT: Robert P. Peery
TITLE OF INVENTION: Methods and Reagents for
TITLE OF INVENTION: Detecting Fungal Pathogens In a
TITLE OF INVENTION: Biological Sample
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 S. Wacker Drive Suite 3200
CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS Word 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/360,606B
FILING DATE: December 21, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Berghoff, Paul H.
REGISTRATION NUMBER: 30,243
REFERENCE/DOCKET NUMBER: 94,319
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312)913-0001
TELEFAX: (312)913-0002
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 1856 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Saccharomyces cerevisiae
US-08-360-6068-29

Query Match 2.7%; Score 16; DB 2; Length 1856;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 gtagtgcgcacgcga 154
|||||
Db 167 gtagtgcgcacgcga 152

RESULT 26

US-08-673-388-9/c
Sequence 9, Application US/08673388
Patent No. 5958745
GENERAL INFORMATION:
APPLICANT: Gruys, Kenneth J.
APPLICANT: Mitsky, Timothy A.

APPLICANT: Kishore, Ganesh M.
APPLICANT: Slater, Steven C.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stark, David M.
APPLICANT: Hinchee, Maud A. W.
APPLICANT: Clemente, Thomas E.
APPLICANT: Connor-Ward, Dannelte V.
APPLICANT: Fedele, Mary J.
APPLICANT: Fry, Joyce E.
APPLICANT: Howe, Arlene R.
APPLICANT: Rozman, Renee J.
TITLE OF INVENTION: Methods of Optimizing Substrate Pools and
TITLE OF INVENTION: Biosynthesis of Poly-B-hydroxybutyrate-co-poly-B-hydroxyval
TITLE OF INVENTION: in Bacteria and Plants
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Charles E. Cohen, Monsanto Company, BBRF
STREET: 700 Chesterfield Parkway No. 5958745th
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/673,388
FILING DATE: 13-MAR-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Cohen, Charles E.
REGISTRATION NUMBER: 34,565
REFERENCE/DOCKET NUMBER: 38-21(13618)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6224
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1864 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-673-388-9

Query Match 2.7%; Score 16; DB 2; Length 1864;
Best Local Similarity 100.0%; Pred. No. 39;
Matches <16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 506 ccagcgcgcgcctcgtc 521
|||||
Db 1073 CCAGCGCGCGCTCGTC 1058

RESULT 27
US-08-614-877-9/c
Sequence 9, Application US/08614877
Patent No. 5959179
GENERAL INFORMATION:
APPLICANT: Gruys, Kenneth J.
APPLICANT: Mitsky, Timothy A.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Slater, Steven C.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stark, David M.
APPLICANT: Hinchee, Maud A. W.
APPLICANT: Clemente, Thomas E.
APPLICANT: Connor-Ward, Dannelte V.
APPLICANT: Fedele, Mary J.
APPLICANT: Fry, Joyce E.

APPLICANT: Howe, Arlene R.
APPLICANT: Rozman, Renee J.
TITLE OF INVENTION: Methods of Optimizing Substrate Pools and
TITLE OF INVENTION: Biosynthesis of Poly-B-hydroxybutyrate-co-poly-B-hydroxyvalerate
TITLE OF INVENTION: In Bacteria and Plants
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Charles E. Cohen, Monsanto Company, BBAF
STREET: 700 Chesterfield Parkway No. 5959179th
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/614,877
FILING DATE: 13-MAR-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Cohen, Charles E.
REGISTRATION NUMBER: 34,565
REFERENCE/DOCKET NUMBER: 38-21(10695)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6224
TELEFAX: (314)537-6047
INFORMATION FOR SEQ. ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1864 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-614-877-9

Query Match 2.7%; Score 16; DB 2; Length 1864;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 506 ccagcgcgctgc 521
|||||

Db 1073 CCAGCGCGCTGCCTC 1058

RESULT 28
US-08-864-038A-1
Sequence 1, Application US/08864038A
Patent No. 6001592
GENERAL INFORMATION:
APPLICANT: Kunio NAKASHIMA et al.
TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR
CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: 812-5 Hirano
STREET: Isshinuden
CITY: Tsu-city
STATE: Mie-prefecture
COUNTRY: JAPAN
ZIP: 514-01
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: Microsoft Windows 95
SOFTWARE: Word Perfect 6.1
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/864,038A
FILING DATE: May 28, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-184459
FILING DATE: 15-July-1996
ATTORNEY/AGENT INFORMATION:
NAME: C. Bruce Hamburg
REGISTRATION NUMBER: 22,389
REFERENCE/DOCKET NUMBER: F-5610
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)986-2340
TELEFAX: (212)953-7733
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2214
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: Pinctada fucata
CELL TYPE: mantle epithelial cell
US-08-864-038A-1

Query Match 2.7%; Score 16; DB 3; Length 2214;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 ctgcgcgtgc 230
|||||

Db 1544 CTGCCTGCTGCATC 1559

RESULT 29
US-09-036-987A-25
Sequence 25, Application US/09036987A
Patent No. 6143526
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Treadway, Patil J.
APPLICANT: Turner, Jan R.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R.
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ. ID NO: 25:

SEQUENCE CHARACTERISTICS:
LENGTH: 2310 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 88..1077
NAME/KEY: CDS
LOCATION: 1165..1992
US-09-036-987A-25

Query Match 2.7%; Score 16; DB 3; Length 2310;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 395 cgaacaggtctacgac 410
|||||
Db 468 CGACGAGGCTACGAC 483

RESULT 30
US-08-864-038A-2
Sequence 2, Application US/08864038A
Patent No. 6001592

GENERAL INFORMATION:
APPLICANT: Kunio NAKASHIMA et al.
TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR
TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: 812-5 Hirano
STREET: Isshinden
CITY: Tsu-city
STATE: Mie-prefecture
COUNTRY: JAPAN
ZIP: 514-01

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
OPERATING SYSTEM: Microsoft Windows 95
SOFTWARE: Word Perfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/864,038A
FILING DATE: May 28, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-184459
FILING DATE: 15-July-1996
ATTORNEY/AGENT INFORMATION:
NAME: C. Bruce Hamburg
REGISTRATION NUMBER: 22,389
REFERENCE/DOCKET NUMBER: F-5610
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)986-2340
TELEFAX: (212)953-7733
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3331
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE: Pinctada fucata
ORGANISM: Pinctada fucata
CELL TYPE: mantle epithelial cell
FEATURE: mRNA
LOCATION: from 1 to 3331

IDENTIFICATION METHOD: E (by experiment)
US-08-864-038A-2

Query Match 2.7%; Score 16; DB 3; Length 3331;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 215 ctgccgctgctgcatc 230
|||||
Db 1593 CTGCCGCTGCTGCATC 1608

RESULT 31
US-08-864-038A-4
Sequence 4, Application US/08864038A
Patent No. 6001592

GENERAL INFORMATION:
APPLICANT: Kunio NAKASHIMA et al.
TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR
TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBOD
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: 812-5 Hirano
STREET: Isshinden
CITY: Tsu-city
STATE: Mie-prefecture
COUNTRY: JAPAN
ZIP: 514-01

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
OPERATING SYSTEM: Microsoft Windows 95
SOFTWARE: Word Perfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/864,038A
FILING DATE: May 28, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-184459
FILING DATE: 15-July-1996
ATTORNEY/AGENT INFORMATION:
NAME: C. Bruce Hamburg
REGISTRATION NUMBER: 22,389
REFERENCE/DOCKET NUMBER: F-5610
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)986-2340
TELEFAX: (212)953-7733
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3331
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE: Pinctada fucata
ORGANISM: Pinctada fucata
CELL TYPE: mantle epithelial cell
FEATURE:
NAME/KEY: CDS
LOCATION: from 50 to 2263
IDENTIFICATION METHOD: P (by similarity to some other pattern)
US-08-864-038A-4

Query Match 2.7%; Score 16; DB 3; Length 3331;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 215 ctgccgctgctgcatc 230
|||||

Db 1593 CTGCGCTGCTGATC 1608

RESULT 32

US-08-854-585-1/c
Sequence 1, Application US/08854585

Patent No. 6114140

GENERAL INFORMATION:

APPLICANT: Tonks, Nicholas K. and stman, Arne

TITLE OF INVENTION: Density Enhanced Protein Tyrosine Phosphatase

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 233 South Wacker Drive, Suite 6300

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/854,585

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/237,940

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Borun, Michael F.

REGISTRATION NUMBER: 25,447

REFERENCE/DOCKET NUMBER: 27866/31954

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300

TELEFAX: 312-474-0448

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 5117 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 350..4364

US-08-854-585-1

Query Match 2.7%; Score 16; DB 3; Length 5117;

Best Local Similarity 100.0%; Pred. No. 39;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 263 agcgagcgcgccctcc 278

Db 382 AGCGAGCGCGCCTCC 367

RESULT 33

PCT-US95-05512-1/c

Sequence 1, Application PC/TUS9505512

GENERAL INFORMATION:

APPLICANT: Tonks, Nicholas K. and stman, Arne

TITLE OF INVENTION: Density Enhanced Protein Tyrosine

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

STREET: 233 South Wacker Drive, Suite 6300

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/05512

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Borun, Michael F.

REGISTRATION NUMBER: 25,447

REFERENCE/DOCKET NUMBER: 27866/31954

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300

TELEFAX: 312-474-0448

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 5117 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 350..4364

PCT-US95-05512-1

Query Match 2.7%; Score 16; DB 5; Length 5117;

Best Local Similarity 100.0%; Pred. No. 39;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 263 agcgagcgcgccctcc 278

Db 382 AGCGAGCGCGCCTCC 367

RESULT 34

PCT-US95-13749-3

Sequence 3, Application PC/TUS9513749

GENERAL INFORMATION:

APPLICANT: Amgen Inc.

TITLE OF INVENTION: COMPOSITIONS FOR INCREASED

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Amgen Inc.

STREET: 1840 Denavilland Drive

CITY: Thousand Oaks

STATE: California

COUNTRY: USA

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/13749

FILING DATE:

CLASSIFICATION:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 6545 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 3630..4820

PCT-US95-13749-3

Query Match 2.7%; Score 16; DB 5; Length 6545;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 gcgacgcgcaacg 160
|||||
DB 5380 GCGACGCGCAACCG 5395

RESULT 35
US-08-056-200-93
; Sequence 93, Application US/08056200
; Patent No. 5616500
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/056,200
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fedrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9551 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1507..1644
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1645..2511
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2512..8070
US-08-056-200-93

Query Match 2.7%; Score 16; DB 1; Length 9551;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 106 agggaagacagcagc 121

DB 6658 AGGGAAGACAGCAGC 6673
|||||

RESULT 36
US-08-800-644-93
; Sequence 93, Application US/08800644
; Patent No. 5958752
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,644
; FILING DATE: 14-FEB-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/056,200
; FILING DATE: 30-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fedrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9551 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1507..1644
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1645..2511
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2512..8070
US-08-800-644-93

Query Match 2.7%; Score 16; DB 2; Length 9551;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 106 agggaagacagcagc 121
|||||
DB 6658 AGGGAAGACAGCAGC 6673

RESULT 37
US-08-633-779-2/C
Sequence 2, Application US/08633779
Patent No. 5723332
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: TRANSLATIONAL ENHANCER DNA
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderhye
STREET: 1100 No. 5723332th Giebe Road, 8th Floor
CITY: Arlington
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/633,779
FILING DATE: 19-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Leonard C. Mitchard
REGISTRATION NUMBER: 29,009
REFERENCE/DOCKET NUMBER: 604-381
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-633-779-2

Query Match 2.5%; Score 15; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 425 ctaccgccccgcagc 439
|||||

Db 29 ctaccgccccgcagc 15

RESULT 38
US-08-238-863-14
Sequence 14, Application US/08238863
Patent No. 5503978
GENERAL INFORMATION:
APPLICANT: SCHNEIDER, D. J., GOLD, L., AND FEIGON, J.
TITLE OF INVENTION: HIGH-AFFINITY ssDNA LIGANDS OF HIV-1
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: Beaton & Swanson, P.C.
STREET: 4582 South Ulster Street Parkway, Suite
STREET: #403
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80237
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/238,863
FILING DATE: 6-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson

APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX17
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 850-9900
FAX: (303) 850-9401
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 81
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-238-863-14

Query Match 2.5%; Score 15; DB 1; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 242 cccctgttcagcgc 256
|||||

Db 30 cccctgttcagcgc 44

RESULT 39
US-08-443-407-14
Sequence 14, Application US/08443407
Patent No. 5786462
GENERAL INFORMATION:
APPLICANT: SCHNEIDER, D. J., GOLD, L., AND FEIGON, J.
TITLE OF INVENTION: HIGH-AFFINITY ssDNA LIGANDS OF HIV-1 REVERSE
TITLE OF INVENTION: LIGANDS OF HIV-1 REVERSE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/443,407
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/238,863
FILING DATE: 6-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson

REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX17/C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 81
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-443-407-14

Query Match 2.5%; Score 15; DB 1; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 242 cccctgttcagcgc 256
|||||
DB 30 CCCCTGTTCAAGCCG 44

RESULT 40
PCT-US95-05600-158
Sequence 158, Application PC/TUS9505600
GENERAL INFORMATION:
APPLICANT: GOLD, LARRY
APPLICANT: NIEWLANDT, DAN
APPLICANT: WECKER, MATTHEW
APPLICANT: SCHNEIDER, DANIEL J.
APPLICANT: FEIGON, JOLI
APPLICANT: ALLEN, PATRICK
APPLICANT: SULENGER, BRUCE A.
APPLICANT: DOUDNA, JENNIFER, A.
TITLE OF INVENTION: HIGH-AFFINITY LIGANDS OF
TITLE OF INVENTION: INSULIN RECEPTOR ANTIBODIES, TACHYKININ SUBSTANCE
TITLE OF INVENTION: P, HIV INTEGRASE AND HIV-1 REVERSE TRANSCRIPTASE
NUMBER OF SEQUENCES: 239
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MG
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05600
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/238,863
FILING DATE: 06-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/248,632
FILING DATE: 24-MAY-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/303,362
FILING DATE: 09-SEPTEMBER-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/361,795
FILING DATE: 21-DECEMBER-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/117,991
FILING DATE: 08-SEPTEMBER-1993
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/931,473
FILING DATE: 17-AUGUST-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX17/PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 158:
SEQUENCE CHARACTERISTICS:
LENGTH: 81 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US95-05600-158

Query Match 2.5%; Score 15; DB 5; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 242 cccctgttcagcgc 256
|||||
DB 30 CCCCTGTTCAAGCCG 44

RESULT 41
US-08-994-946A-1
Sequence 1, Application US/08994946A
Patent No. 6046317
GENERAL INFORMATION:
APPLICANT: Koulou, Markku
APPLICANT: Karvonen, Matti
APPLICANT: Pesonen, Ullamari
TITLE OF INVENTION: A DNA Molecule Encoding a Mutant
TITLE OF INVENTION: Prepro-Neuropeptide Y, a Mutant Signal Peptide, and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
STREET: 555 13th Street NW, Suite 701-E
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/994,946A
FILING DATE: 19-DEC-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 2328-110
TELECOMMUNICATION INFORMATION:

```

? TELEPHONE: 202-783-6040
? TELEFAX: 202-783-6031
? INFORMATION FOR SEQ ID NO: 1:
?
? SEQUENCE CHARACTERISTICS:
?
? LENGTH: 325 base pairs
?
? TYPE: nucleic acid
?
? STRANDEDNESS: double
?
? TOPOLOGY: linear
?
? MOLECULE TYPE: DNA (genomic)
US-08-994-946A-1

```

| | | | | |
|-----------------------|-----------------|---------------|-----------|-------------|
| Query Match | 2.5% | Score 15: | DB 3: | Length 325; |
| Best Local Similarity | 100.0% | Pred. No. | 1.3e+02; | |
| Matches 15; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |

| | | | |
|----|-----|----------------|-----|
| Oy | 274 | cctcctgcgcgcg | 288 |
| | | | |
| Db | 155 | CCTCCTGCCGCGCG | 169 |

```

RESULT 42
US-09-123-492A-3/c
: Sequence 3, Application US/09123492A
: Patent No. 6197640
: GENERAL INFORMATION:
: APPLICANT: KLINEFELTER, Gary
: TITLE OF INVENTION: METHOD FOR EVALUATING AND AFFECTING MALE FERTILITY
: FILE REFERENCE: KLINEFELTER-1B
: CURRENT APPLICATION NUMBER: US/09/123,492A
: CURRENT FILING DATE: 1998-07-28
: PRIOR APPLICATION NUMBER: 08/593,677
: PRIOR FILING DATE: 1997-01-29
: PRIOR APPLICATION NUMBER: PCT/US97/01725
: PRIOR FILING DATE: 1997-01-29
: PRIOR APPLICATION NUMBER: 60/082,753
: PRIOR FILING DATE: 1998-04-23
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 3
: LENGTH: 837
: TYPE: DNA
: ORGANISM: Rat sperm
: FEATURE:
: OTHER INFORMATION: Coding region from nucleotides 52-618 encodes for
: OTHER INFORMATION: residues 14-202 of SEQ ID NO:5
: US-09-123-492A-3

```

| | | | | |
|--------------------------|--------|------------|---------|------------|
| Query Match | 2.5% | Score 15 | DB 4 | Length 837 |
| Best Local Similarity | 100.0% | Pred. No. | 1.3e+02 | |
| Matches 15, Conservative | 0 | Mismatches | 0 | Gaps 0 |

```

QY      265  gcaagccgcgcctcct 279
          ||| ||| ||| ||| |||
Db      32  GCAGCCGCGCCTCT 18

```

RESULT 43
 US-08-633-779-1/c
 : Sequence 1, Application US/08633779
 : Patent No. 5723332
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT:
 :
 : TITLE OF INVENTION: TRANSLATIONAL ENHANCER DNA
 :
 : NUMBER OF SEQUENCES: 8
 :
 : CORRESPONDENCE ADDRESS:
 :
 : ADDRESSEE: Nixon & Vanderhpe
 : STREET: 1100 NO. 5723332th Glebe Road, 8th Floor
 : CITY: Arlington
 : COUNTRY: U.S.A.
 : ZIP: 22201-4714
 :
 : COMPUTER READABLE FORM:

1 MEDIUM TYPE: Floppy disk
 2 COMPUTER: IBM PC compatible
 3 OPERATING SYSTEM: PC-DOS/MS-DOS
 4 SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
 5 CURRENT APPLICATION DATA:
 6 APPLICATION NUMBER: US/08/633,779
 7 FILING DATE: 19-Apr-1996
 8 CLASSIFICATION: 435
 9 ATTORNEY/AGENT INFORMATION:
 10 NAME: Leonard C. Mltchard
 11 REGISTRATION NUMBER: 29,009
 12 REFERENCE/DOCKET NUMBER: 604-381
 13 INFORMATION FOR SEQ ID NO: 1:
 14 SEQUENCE CHARACTERISTICS:
 15 LENGTH: 838 base pairs
 16 TYPE: nucleic acid
 17 STRANDEDNESS: double
 18 TOPOLOGY: linear
 19 MOLECULE TYPE: DNA (genomic)
 20 US-08-633-779-1

```

Query Match Similarity 2.5%; Score 15; DB 1; Length 838;
Best Local Similarity 100.0%; Pred. No. 1 3e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

| QY | 425 | ctaccgccccgagc | 439 |
|----|-----|----------------|-----|
| | | | |
| Db | 407 | CTACCGCCCGCAGC | 393 |

RESULT 44
US-09-154-874-1/c
Sequence 1, Application US/09154874
Patent No. 6054636
GENERAL INFORMATION:
APPLICANT: FADER, GARY MICHAEL
TITLE OF INVENTION: ISOFALVONE BIOSYNTHETIC ENZYMES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/154,874
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/931,668
FILING DATE: SEPTEMBER 17, 1997
ATTORNEY/AGENT INFORMATION:
NAME: MAJARIAN, WILLIAM R.
REGISTRATION NUMBER: 41,173
REFERENCE/DOCKET NUMBER: BB-1098-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-4926
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 846 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:

NAME/KEY: CDS
LOCATION: 40..693
US-09-154-874-1

Query Match 2.5%; Score 15; DB 3; Length 846;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 ttccagccagccca 263
|||||
DB 351 TTCAGCCAGCCCA 337

RESULT 45
US-09-082-092-8/c

Sequence 8, Application US/09082092
Patent No. 6251628

GENERAL INFORMATION:

APPLICANT: Nakao, Atsuhito

APPLICANT: Moren, Anita

APPLICANT: Heuchel, Rainer

APPLICANT: Itoh, Susumu

APPLICANT: Afirakhte, Mozhgan

APPLICANT: Souchehytskyi, Serhiy

APPLICANT: Brodin, Greger

APPLICANT: Landstrom, Marene

APPLICANT: Heldin, Nils-Erik

APPLICANT: Heldin, Carl-Henrik

TITLE OF INVENTION: SMAD7 AND USES THEREOF

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wolf, Greenfield & Sacks, P.C.

STREET: 600 Atlantic Avenue

CITY: Boston

STATE: MA

COUNTRY: U.S.A.

ZIP: 02210-2211

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/082.092

FILING DATE: 20-MAY-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/047,221

FILING DATE: 20-MAY-1997

APPLICATION NUMBER: 60/060,465

FILING DATE: 30-SEP-1997

APPLICATION NUMBER: 60/075,940

FILING DATE: 25-FEB-1998

APPLICATION NUMBER: 60/077,033

FILING DATE: 06-MAR-1998

ATTORNEY/AGENT INFORMATION:

NAME: Van Amsterdam, John R.

REGISTRATION NUMBER: 40,212

REFERENCE/DOCKET NUMBER: L0461/7032

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-720-3500

TELEFAX: 617-720-2441

TELEX:

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 1281 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-09-082-092-8

Query Match 2.5%; Score 15; DB 4; Length 1281;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 ccccccgttcagcg 255
|||||
DB 710 CCCCTGTTTCAGCG 696

RESULT 46
US-08-176-427B-7/c

Sequence 7, Application US/08176427B

Patent No. 5789543

GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.

APPLICANT: McMahon, Andrew P.

APPLICANT: Tablin, Clifford J.

TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing

TITLE OF INVENTION: Proteins and Uses Related Thereto

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII(text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/176.427B

FILING DATE: 30-DEC-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: HMI-006

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1313 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1314

US-08-176-427B-7

Query Match 2.5%; Score 15; DB 1; Length 1313;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 502 ggttcagcgagcgcc 516
|||||
DB 1133 GGTGCCAGCGCGCC 1119

RESULT 47

US-08-356-060A-4/c

Sequence 4, Application US/08356060A

Patent No. 5844079

GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.


```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/855,910
; FILING DATE: 14-MAY-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: Brook, David E.
;   REGISTRATION NUMBER: 22,592
;   REFERENCE/DOCKET NUMBER: CP195-08
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (781) 861-6240
;   TELEFAX: (781) 861-9540
; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1338 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 187..1194
;   US-08-855-910-5

Query Match      2.5%; Score 15; DB 4; Length 1338;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 atcgtaacgacgacg 43
      |||
Db 1259 ATCGTACAGCCGACG 1273
```

```

RESULT 50
US-08-651-136C-9
; Sequence 9, Application US/08651136C
; Patent No. 6001639
; GENERAL INFORMATION:
;   APPLICANT: Schulein, Martin
;   APPLICANT: Andersen, Lene N.
;   APPLICANT: Lassen, Soren F.
;   APPLICANT: Kauppinen, Markus S.
;   APPLICANT: Lange, Lene
;   APPLICANT: Nielsen, Rudy I.
;   APPLICANT: Ihara, Michiko
;   APPLICANT: Takagi, Shinobu
;   TITLE OF INVENTION: No. 6001639e1 Endoglucanases
;   NUMBER OF SEQUENCES: 109
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: No. 6001639o No. 6001639disk of No. 6001639th America, Inc.
;   STREET: 405 Lexington Avenue, 64th Floor
;   CITY: New York
;   STATE: New York
;   COUNTRY: United States of America
;   ZIP: 10174-6401
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/651,136C
;   FILING DATE: 21-MAY-1996
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Lambiris, Elias J.
;     REGISTRATION NUMBER: 33,728
;     REFERENCE/DOCKET NUMBER: 4366,200-US
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 212-867-0123
;     TELEFAX: 212-878-9655
;   INFORMATION FOR SEQ ID NO: 9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 1423 base pairs
```

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 110..1156
;   US-08-651-136C-9

Query Match      2.5%; Score 15; DB 3; Length 1423;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 gccggccaacgacg 401
      |||
Db 979 GCCGGCCACGACGAG 993
```

Search completed: September 21, 2001, 16:24:12
Job time: 8823 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 13:40:55 ; Search time 2753.51 Seconds
(without alignments)
3187.722 Million cell updates/sec

Title: US-09-138-735-1_COPY_1232_1825

Perfect score: 594
Sequence: 1 caggtacagcgtacgctt.....aaacgagctccgtgagcgtata 594

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 14155048 seqs, 7388405095 residues

Word size: 0

Total number of hits satisfying chosen parameters: 28310096

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database:

Pending_Patents_MA_Main.*
1: /cgnl_7/ptodata/1/pna/PCrus.COMB.seq.*
2: /cgnl_7/ptodata/1/pna/US06.COMB.seq.*
3: /cgnl_7/ptodata/1/pna/US07.COMB.seq.*
4: /cgnl_7/ptodata/1/pna/US08.COMB.seq.*
5: /cgnl_7/ptodata/1/pna/US081.COMB.seq.*
6: /cgnl_7/ptodata/1/pna/US082.COMB.seq.*
7: /cgnl_7/ptodata/1/pna/US083.COMB.seq.*
8: /cgnl_7/ptodata/1/pna/US084.COMB.seq.*
9: /cgnl_7/ptodata/1/pna/US085.COMB.seq.*
10: /cgnl_7/ptodata/1/pna/US086.COMB.seq.*
11: /cgnl_7/ptodata/1/pna/US087.COMB.seq.*
12: /cgnl_7/ptodata/1/pna/US088.COMB.seq.*
13: /cgnl_7/ptodata/1/pna/US089.COMB.seq.*
14: /cgnl_7/ptodata/1/pna/US090.COMB.seq.*
15: /cgnl_7/ptodata/1/pna/US091.COMB.seq.*
16: /cgnl_7/ptodata/1/pna/US092.COMB.seq.*
17: /cgnl_7/ptodata/1/pna/US093.COMB.seq.*
18: /cgnl_7/ptodata/1/pna/US094.COMB.seq.*
19: /cgnl_7/ptodata/1/pna/US095.COMB.seq.*
20: /cgnl_7/ptodata/1/pna/US096.COMB.seq.*
21: /cgnl_7/ptodata/1/pna/US097.COMB.seq.*
22: /cgnl_7/ptodata/1/pna/US098.COMB.seq.*
23: /cgnl_7/ptodata/1/pna/US099.COMB.seq.*
24: /cgnl_7/ptodata/1/pna/US100.COMB.seq.*
25: /cgnl_7/ptodata/1/pna/US101.COMB.seq.*
26: /cgnl_7/ptodata/1/pna/US102.COMB.seq.*
27: /cgnl_7/ptodata/1/pna/US103.COMB.seq.*
28: /cgnl_7/ptodata/1/pna/US104.COMB.seq.*
29: /cgnl_7/ptodata/1/pna/US105.COMB.seq.*
30: /cgnl_7/ptodata/1/pna/US106.COMB.seq.*
31: /cgnl_7/ptodata/1/pna/US107.COMB.seq.*
32: /cgnl_7/ptodata/1/pna/US108.COMB.seq.*
33: /cgnl_7/ptodata/1/pna/US109.COMB.seq.*
34: /cgnl_7/ptodata/1/pna/US110.COMB.seq.*
35: /cgnl_7/ptodata/1/pna/US111.COMB.seq.*
36: /cgnl_7/ptodata/1/pna/US112.COMB.seq.*
37: /cgnl_7/ptodata/1/pna/US113.COMB.seq.*
38: /cgnl_7/ptodata/1/pna/US114.COMB.seq.*
39: /cgnl_7/ptodata/1/pna/US115.COMB.seq.*
40: /cgnl_7/ptodata/1/pna/US116.COMB.seq.*
41: /cgnl_7/ptodata/1/pna/US117.COMB.seq.*
42: /cgnl_7/ptodata/1/pna/US118.COMB.seq.*
43: /cgnl_7/ptodata/1/pna/US119.COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1 | 594 | 100.0 | 3402 | US-08-988-242-1 | Sequence 1, Appl1 |
| 2 | 594 | 100.0 | 3402 | US-09-138-736-1 | Sequence 1, Appl1 |
| 3 | 22 | 3.7 | 22 | US-08-988-242-5 | Sequence 5, Appl1 |
| 4 | 22 | 3.7 | 22 | US-09-138-736-9 | Sequence 9, Appl1 |
| 5 | 20 | 3.4 | 718 | US-09-654-617-126156 | Sequence 126156, |
| 6 | 20 | 3.4 | 718 | US-09-684-016-126156 | Sequence 126156, |
| 7 | 20 | 3.4 | 5679 | US-09-620-392-64793 | Sequence 64793, A |
| 8 | 20 | 3.4 | 5679 | US-09-702-134-21010 | Sequence 21010, A |
| 9 | 20 | 3.4 | 11335 | US-06-173-464-11335 | Sequence 11335, A |
| 10 | 20 | 3.4 | 11335 | US-06-191-637-11335 | Sequence 11335, A |
| 11 | 20 | 3.4 | 11335 | US-06-191-637-11335 | Sequence 11335, A |
| 12 | 20 | 3.4 | 11338 | US-06-167-217-139986 | Sequence 10986, A |
| 13 | 20 | 3.4 | 91470 | US-09-534-859-19 | Sequence 13998, A |
| 14 | 20 | 3.4 | 226706 | US-09-528-237A-1285 | Sequence 19, Appl1 |
| 15 | 19 | 3.2 | 718 | US-06-182-316-2055 | Sequence 1285, Ap |
| 16 | 19 | 3.2 | 9855 | US-09-620-392-32691 | Sequence 2055, Ap |
| 17 | 19 | 3.2 | 9855 | US-09-620-392-67713 | Sequence 32691, A |
| 18 | 19 | 3.2 | 11873 | US-09-407-562-31 | Sequence 67713, A |
| 19 | 19 | 3.2 | 11878 | US-09-407-562-31 | Sequence 32, Appl1 |
| 20 | 19 | 3.2 | 11883 | US-09-407-562-28 | Sequence 31, Appl1 |
| 21 | 19 | 3.2 | 15505 | US-09-620-392-10632 | Sequence 28, Appl1 |
| 22 | 19 | 3.2 | 25670 | US-09-620-392-4251 | Sequence 10632, A |
| 23 | 19 | 3.2 | 31556 | US-09-702-134-5295 | Sequence 4251, Ap |
| 24 | 19 | 3.2 | 159989 | US-09-702-134-5301 | Sequence 5295, Ap |
| 25 | 18 | 3.0 | 18 | US-08-988-242-4 | Sequence 5301, Ap |
| 26 | 18 | 3.0 | 18 | US-09-138-736-5 | Sequence 4, Appl1 |
| 27 | 18 | 3.0 | 18 | US-09-138-736-8 | Sequence 5, Appl1 |
| 28 | 18 | 3.0 | 70 | US-09-366-691A-1504 | Sequence 8, Appl1 |
| 29 | 18 | 3.0 | 85 | US-09-076-667-928 | Sequence 1504, Ap |
| 30 | 18 | 3.0 | 85 | US-09-540-212A-11685 | Sequence 928, Ap |
| 31 | 18 | 3.0 | 85 | US-09-540-212A-11685 | Sequence 15688, A |
| 32 | 18 | 3.0 | 180 | US-06-048-002-928 | Sequence 928, App |
| 33 | 18 | 3.0 | 180 | US-09-513-991-2039 | Sequence 2039, App |
| 34 | 18 | 3.0 | 234 | US-08-810-326-2151 | Sequence 2039, App |
| 35 | 18 | 3.0 | 234 | US-09-540-212A-11685 | Sequence 2151, Ap |
| 36 | 18 | 3.0 | 237 | US-08-878-507-222 | Sequence 11685, A |
| 37 | 18 | 3.0 | 237 | US-09-540-212A-11685 | Sequence 222, App |
| 38 | 18 | 3.0 | 237 | US-06-016-092-222 | Sequence 16821, A |
| 39 | 18 | 3.0 | 248 | US-09-100-454-2483 | Sequence 222, App |
| 40 | 18 | 3.0 | 248 | US-09-540-212A-35497 | Sequence 2483, Ap |
| 41 | 18 | 3.0 | 249 | US-06-051-749-2483 | Sequence 35497, A |
| 42 | 18 | 3.0 | 249 | US-06-048-722-3053 | Sequence 2483, Ap |
| 43 | 18 | 3.0 | 253 | US-08-992-625-3252 | Sequence 3053, Ap |
| 44 | 18 | 3.0 | 253 | US-09-540-764-1048 | Sequence 3252, Ap |
| 45 | 18 | 3.0 | 261 | US-06-197-872-21718 | Sequence 1048, A |
| 46 | 18 | 3.0 | 261 | US-08-104-507A-4311 | Sequence 21718, A |
| | | | 5 | US-08-104-507C-4311 | Sequence 4311, Ap |

| | | | | | | | | | | | | | |
|-------|----|-----|------|----|----------------------|---------------------|-------|----|-----|--------|----|----------------------|---------------------|
| C 47 | 18 | 3.0 | 269 | 16 | US-09-270-849B-53924 | Sequence 53924, A | C 120 | 18 | 3.0 | 1589 | 25 | US-09-641-377-783 | Sequence 783, App |
| C 48 | 18 | 3.0 | 271 | 16 | US-09-298-329A-4053 | Sequence 4053, Ap | C 121 | 18 | 3.0 | 1707 | 25 | US-09-644-867-5948 | Sequence 5948, Ap |
| C 49 | 18 | 3.0 | 271 | 40 | US-60-087-762-211 | Sequence 211, App | C 122 | 18 | 3.0 | 1707 | 25 | US-09-652-914-7995 | Sequence 7995, Ap |
| C 50 | 18 | 3.0 | 271 | 41 | US-60-092-036-941 | Sequence 941, App | C 123 | 18 | 3.0 | 1896 | 22 | US-09-586-939-130 | Sequence 130, App |
| C 51 | 18 | 3.0 | 284 | 48 | US-60-164-443-401 | Sequence 401, App | C 124 | 18 | 3.0 | 2218 | 49 | US-60-172-373-22602 | Sequence 22602, A |
| C 52 | 18 | 3.0 | 290 | 12 | US-08-826-438-125 | Sequence 125, App | C 125 | 18 | 3.0 | 2520 | 7 | US-08-340-426B-50 | Sequence 50, Appl |
| C 53 | 18 | 3.0 | 290 | 12 | US-08-951-197-125 | Sequence 125, App | C 126 | 18 | 3.0 | 2520 | 8 | US-08-450-673A-50 | Sequence 50, Appl |
| C 54 | 18 | 3.0 | 290 | 20 | US-09-539-806-21607 | Sequence 21607, A | C 127 | 18 | 3.0 | 2520 | 8 | US-08-450-673A-50 | Sequence 50, Appl |
| C 55 | 18 | 3.0 | 290 | 33 | US-60-013-696-4335 | Sequence 4335, Ap | C 128 | 18 | 3.0 | 2762 | 60 | US-60-281-593-39 | Sequence 39, Appl |
| C 56 | 18 | 3.0 | 297 | 16 | US-09-266-640A-3874 | Sequence 3874, Ap | C 129 | 18 | 3.0 | 2803 | 25 | US-09-652-814-979 | Sequence 979, App |
| C 57 | 18 | 3.0 | 297 | 16 | US-09-266-640B-3874 | Sequence 3874, Ap | C 130 | 18 | 3.0 | 2803 | 29 | US-09-726-172-11733 | Sequence 11733, Ap |
| C 58 | 18 | 3.0 | 297 | 17 | US-09-313-294A-6114 | Sequence 6114, Ap | C 131 | 18 | 3.0 | 2803 | 29 | US-09-726-172-11733 | Sequence 11733, Ap |
| C 59 | 18 | 3.0 | 297 | 40 | US-60-086-722-6114 | Sequence 6114, Ap | C 132 | 18 | 3.0 | 2810 | 17 | US-09-397-424-5114 | Sequence 5114, Ap |
| C 60 | 18 | 3.0 | 297 | 40 | US-60-087-762-208 | Sequence 208, App | C 133 | 18 | 3.0 | 2810 | 17 | US-09-397-424-5114 | Sequence 5114, Ap |
| C 61 | 18 | 3.0 | 297 | 41 | US-60-092-036-939 | Sequence 939, App | C 134 | 18 | 3.0 | 2817 | 17 | US-09-338-663-1298 | Sequence 1298, Ap |
| C 62 | 18 | 3.0 | 298 | 17 | US-09-304-517A-5225 | Sequence 5225, Ap | C 135 | 18 | 3.0 | 2817 | 25 | US-09-652-814-979 | Sequence 979, App |
| C 63 | 18 | 3.0 | 298 | 17 | US-09-371-146A-5225 | Sequence 5225, Ap | C 136 | 18 | 3.0 | 2817 | 25 | US-09-652-814-979 | Sequence 979, App |
| C 64 | 18 | 3.0 | 298 | 25 | US-09-654-617-301849 | Sequence 301849, Ap | C 137 | 18 | 3.0 | 3053 | 53 | US-60-213-355-1146 | Sequence 1146, Ap |
| C 65 | 18 | 3.0 | 298 | 27 | US-09-684-016-301849 | Sequence 301849, Ap | C 138 | 18 | 3.0 | 3078 | 18 | US-09-428-151A-8869 | Sequence 8869, Ap |
| C 66 | 18 | 3.0 | 322 | 26 | US-60-202-214-5107 | Sequence 5107, Ap | C 139 | 18 | 3.0 | 3083 | 19 | US-09-522-303-1419 | Sequence 1419, Ap |
| C 67 | 18 | 3.0 | 322 | 26 | US-09-663-779-4257 | Sequence 4257, Ap | C 140 | 18 | 3.0 | 3143 | 25 | US-60-212-659-815 | Sequence 815, App |
| C 68 | 18 | 3.0 | 328 | 1 | PCT-US01-01329-161 | Sequence 161, App | C 141 | 18 | 3.0 | 3175 | 53 | US-60-212-659-815 | Sequence 815, App |
| C 69 | 18 | 3.0 | 328 | 1 | PCT-US01-01339-1027 | Sequence 1027, Ap | C 142 | 18 | 3.0 | 3175 | 53 | US-60-212-659-815 | Sequence 815, App |
| C 70 | 18 | 3.0 | 339 | 21 | US-09-540-229-178511 | Sequence 178511, Ap | C 143 | 18 | 3.0 | 7959 | 28 | US-09-703-708-635 | Sequence 635, App |
| C 71 | 18 | 3.0 | 348 | 11 | US-08-727-898-198 | Sequence 198, App | C 144 | 18 | 3.0 | 7959 | 48 | US-60-164-320-635 | Sequence 635, App |
| C 72 | 18 | 3.0 | 348 | 32 | US-60-005-018-198 | Sequence 198, App | C 145 | 18 | 3.0 | 7959 | 50 | US-60-164-320-635 | Sequence 635, App |
| C 73 | 18 | 3.0 | 367 | 17 | US-09-304-517A-82813 | Sequence 82813, A | C 146 | 18 | 3.0 | 9227 | 56 | US-60-248-798-120 | Sequence 120, App |
| C 74 | 18 | 3.0 | 367 | 17 | US-09-371-146A-82813 | Sequence 82813, A | C 147 | 18 | 3.0 | 9227 | 56 | US-60-248-798-120 | Sequence 120, App |
| C 75 | 18 | 3.0 | 367 | 17 | US-09-394-745-7320 | Sequence 7320, Ap | C 148 | 18 | 3.0 | 15270 | 1 | PCT-US01-01354-28829 | Sequence 28829, A |
| C 76 | 18 | 3.0 | 367 | 22 | US-09-565-306-17581 | Sequence 17581, A | C 149 | 18 | 3.0 | 15270 | 1 | PCT-US01-01354-28829 | Sequence 28829, A |
| C 77 | 18 | 3.0 | 371 | 47 | US-60-155-006-4578 | Sequence 4578, Ap | C 150 | 18 | 3.0 | 20671 | 53 | US-60-212-656-217 | Sequence 217, App |
| C 78 | 18 | 3.0 | 386 | 51 | US-60-197-872-6873 | Sequence 6873, Ap | C 151 | 18 | 3.0 | 23620 | 55 | US-60-230-435-725 | Sequence 725, App |
| C 79 | 18 | 3.0 | 393 | 51 | US-60-197-872-6873 | Sequence 6873, Ap | C 152 | 18 | 3.0 | 30885 | 40 | US-60-082-302-745 | Sequence 745, App |
| C 80 | 18 | 3.0 | 438 | 1 | PCT-US01-04942A-529 | Sequence 529, App | C 153 | 18 | 3.0 | 32768 | 55 | US-60-230-435-725 | Sequence 725, App |
| C 81 | 18 | 3.0 | 438 | 19 | US-09-519-705-2501 | Sequence 2501, Ap | C 154 | 18 | 3.0 | 33042 | 56 | US-60-245-221-47 | Sequence 47, Appl |
| C 82 | 18 | 3.0 | 438 | 22 | US-09-574-454-2501 | Sequence 2501, Ap | C 155 | 18 | 3.0 | 46405 | 28 | US-09-703-708-728 | Sequence 728, App |
| C 83 | 18 | 3.0 | 442 | 19 | US-09-515-126-5771 | Sequence 5771, Ap | C 156 | 18 | 3.0 | 46405 | 48 | US-60-164-320-728 | Sequence 728, App |
| C 84 | 18 | 3.0 | 442 | 22 | US-09-577-409-5771 | Sequence 5771, Ap | C 157 | 18 | 3.0 | 68750 | 50 | US-60-183-791-728 | Sequence 728, App |
| C 85 | 18 | 3.0 | 444 | 30 | US-09-170-175-4591 | Sequence 4591, Ap | C 158 | 18 | 3.0 | 68750 | 22 | US-09-567-899-1 | Sequence 1, Appl |
| C 86 | 18 | 3.0 | 454 | 20 | US-09-533-559-1832 | Sequence 1832, Ap | C 159 | 18 | 3.0 | 68750 | 22 | US-09-567-899-1 | Sequence 1, Appl |
| C 87 | 18 | 3.0 | 458 | 18 | US-09-401-645-3287 | Sequence 3287, Ap | C 160 | 18 | 3.0 | 68750 | 22 | US-09-568-102-1 | Sequence 1, Appl |
| C 88 | 18 | 3.0 | 459 | 19 | US-09-528-409-105673 | Sequence 105673, Ap | C 161 | 18 | 3.0 | 68750 | 22 | US-09-568-102-1 | Sequence 1, Appl |
| C 89 | 18 | 3.0 | 468 | 1 | PCT-US01-00063-60 | Sequence 60, Appl | C 162 | 18 | 3.0 | 68750 | 22 | US-09-568-102-1 | Sequence 1, Appl |
| C 90 | 18 | 3.0 | 468 | 55 | US-60-236-359-64 | Sequence 64, Appl | C 163 | 18 | 3.0 | 68750 | 22 | US-09-568-102-1 | Sequence 1, Appl |
| C 91 | 18 | 3.0 | 473 | 18 | US-09-489-036-926 | Sequence 926, Appl | C 164 | 18 | 3.0 | 68750 | 43 | US-60-118-906-1 | Sequence 1, Appl |
| C 92 | 18 | 3.0 | 475 | 16 | US-09-240-371-7431 | Sequence 7431, Ap | C 165 | 18 | 3.0 | 71989 | 18 | US-09-443-501A-2 | Sequence 2, Appl |
| C 93 | 18 | 3.0 | 475 | 16 | US-09-293-972-5995 | Sequence 5995, Ap | C 166 | 18 | 3.0 | 71989 | 18 | US-09-443-501A-2 | Sequence 2, Appl |
| C 94 | 18 | 3.0 | 476 | 16 | US-09-234-611-19378 | Sequence 19378, A | C 167 | 18 | 3.0 | 164677 | 54 | US-60-226-176-201 | Sequence 201, App |
| C 95 | 18 | 3.0 | 476 | 16 | US-09-277-227-11623 | Sequence 11623, A | C 168 | 18 | 3.0 | 164677 | 55 | US-60-233-468-201 | Sequence 201, App |
| C 96 | 18 | 3.0 | 478 | 1 | PCT-US01-00063-13298 | Sequence 13298, A | C 169 | 18 | 3.0 | 164677 | 55 | US-60-233-468-201 | Sequence 201, App |
| C 97 | 18 | 3.0 | 478 | 55 | US-60-236-359-7474 | Sequence 7474, Ap | C 170 | 18 | 3.0 | 178888 | 56 | US-60-248-798-28 | Sequence 28, Appl |
| C 98 | 18 | 3.0 | 484 | 23 | US-09-606-776-3048 | Sequence 3048, Ap | C 171 | 18 | 3.0 | 115 | 14 | US-09-078-611-1204 | Sequence 1204, Ap |
| C 99 | 18 | 3.0 | 484 | 25 | US-09-644-867-3664 | Sequence 3664, Ap | C 172 | 18 | 2.9 | 154 | 14 | US-09-016-884-1388 | Sequence 1388, Ap |
| C 100 | 18 | 3.0 | 506 | 6 | US-08-276-163A-13709 | Sequence 13709, A | C 173 | 18 | 2.9 | 154 | 35 | US-60-037-043-1388 | Sequence 1388, Ap |
| C 101 | 18 | 3.0 | 506 | 6 | US-08-276-163B-13709 | Sequence 13709, A | C 174 | 18 | 2.9 | 158 | 16 | US-09-231-925-1568 | Sequence 1568, Ap |
| C 102 | 18 | 3.0 | 506 | 6 | US-08-276-163D-13709 | Sequence 13709, A | C 175 | 18 | 2.9 | 158 | 21 | US-09-540-229-142669 | Sequence 142669, Ap |
| C 103 | 18 | 3.0 | 508 | 25 | US-09-652-914-2822 | Sequence 2822, Ap | C 176 | 18 | 2.9 | 163 | 16 | US-09-288-778A-4508 | Sequence 4508, Ap |
| C 104 | 18 | 3.0 | 538 | 19 | US-09-522-305-313 | Sequence 313, App | C 177 | 18 | 2.9 | 163 | 16 | US-60-089-763-4508 | Sequence 4508, Ap |
| C 105 | 18 | 3.0 | 580 | 25 | US-09-654-617-387081 | Sequence 387081, Ap | C 178 | 18 | 2.9 | 166 | 12 | US-08-867-016-1901 | Sequence 190, App |
| C 106 | 18 | 3.0 | 580 | 27 | US-09-684-016-387081 | Sequence 387081, Ap | C 179 | 18 | 2.9 | 166 | 21 | US-09-540-208-11694 | Sequence 11694, A |
| C 107 | 18 | 3.0 | 605 | 25 | US-09-652-915-1811 | Sequence 1811, Ap | C 180 | 18 | 2.9 | 176 | 18 | US-09-654-617-228211 | Sequence 228211, A |
| C 108 | 18 | 3.0 | 664 | 25 | US-09-654-617-274884 | Sequence 274884, Ap | C 181 | 18 | 2.9 | 176 | 20 | US-09-452-747-574 | Sequence 574, App |
| C 109 | 18 | 3.0 | 664 | 27 | US-09-684-016-274884 | Sequence 274884, Ap | C 182 | 18 | 2.9 | 176 | 43 | US-60-111-910-574 | Sequence 574, App |
| C 110 | 18 | 3.0 | 726 | 16 | US-09-452-991A-9311 | Sequence 9311, Ap | C 183 | 18 | 2.9 | 178 | 13 | US-08-901-904-4790 | Sequence 4790, Ap |
| C 111 | 18 | 3.0 | 730 | 40 | US-60-082-300-3133 | Sequence 3133, Ap | C 184 | 18 | 2.9 | 184 | 21 | US-09-540-208-11694 | Sequence 11694, A |
| C 112 | 18 | 3.0 | 771 | 17 | US-09-364-310-1396 | Sequence 1396, Ap | C 185 | 18 | 2.9 | 197 | 24 | US-09-637-086A-44437 | Sequence 44437, A |
| C 113 | 18 | 3.0 | 891 | 18 | US-09-417-507-438 | Sequence 438, App | C 186 | 18 | 2.9 | 197 | 25 | US-09-654-617-228211 | Sequence 228211, A |
| C 114 | 18 | 3.0 | 940 | 23 | US-09-606-776-5253 | Sequence 5253, Ap | C 187 | 18 | 2.9 | 197 | 27 | US-09-684-016-228211 | Sequence 228211, A |
| C 115 | 18 | 3.0 | 940 | 30 | US-09-770-175-8293 | Sequence 8293, Ap | C 188 | 18 | 2.9 | 198 | 15 | US-09-081-708-348 | Sequence 348, App |
| C 116 | 18 | 3.0 | 1302 | 1 | PCT-US00-06823-54 | Sequence 54, Appl | C 189 | 18 | 2.9 | 198 | 15 | US-09-181-375-87 | Sequence 87, Appl |
| C 117 | 18 | 3.0 | 1353 | 18 | US-09-471-275-9469 | Sequence 9469, Ap | C 190 | 18 | 2.9 | 200 | 14 | US-60-021-275-239 | Sequence 239, App |
| C 118 | 18 | 3.0 | 1380 | 16 | US-09-252-991A-9268 | Sequence 9268, Ap | C 191 | 18 | 2.9 | 200 | 12 | US-08-883-626-607 | Sequence 607, App |
| C 119 | 18 | 3.0 | 1390 | 27 | US-09-698-010-14025 | Sequence 14025, A | C 192 | 18 | 2.9 | 200 | 21 | US-09-540-499-4241 | Sequence 4241, Ap |

| | | | | | | | | | | | | | |
|-----|----|-----|-----|----|----------------------|---------------------|-----|----|-----|-----|----|-----------------------|--------------------|
| 193 | 17 | 2.9 | 200 | 34 | US-60-021-995-607 | Sequence 607, App | 266 | 17 | 2.9 | 246 | 32 | US-60-004-674-614 | Sequence 614, App |
| 194 | 17 | 2.9 | 206 | 12 | US-08-878-504-1667 | Sequence 1667, Ap | 267 | 17 | 2.9 | 246 | 35 | US-60-039-416-48 | Sequence 48, App |
| 195 | 17 | 2.9 | 206 | 21 | US-09-540-229-64908 | Sequence 64908, A | 268 | 17 | 2.9 | 247 | 13 | US-08-903-556-324 | Sequence 324, App |
| 196 | 17 | 2.9 | 206 | 34 | US-60-020-310-1667 | Sequence 1667, Ap | 269 | 17 | 2.9 | 247 | 21 | US-09-540-764-10605 | Sequence 10605, A |
| 197 | 17 | 2.9 | 210 | 13 | US-08-878-669-1858 | Sequence 1858, Ap | 270 | 17 | 2.9 | 247 | 21 | US-60-040-957-83 | Sequence 83, App |
| 198 | 17 | 2.9 | 211 | 13 | US-08-903-802-480 | Sequence 480, App | 271 | 17 | 2.9 | 248 | 20 | US-09-539-806-36947 | Sequence 36947, A |
| 199 | 17 | 2.9 | 212 | 13 | US-08-903-469-4342 | Sequence 4342, Ap | 272 | 17 | 2.9 | 250 | 14 | US-09-010-765-731 | Sequence 731, App |
| 200 | 17 | 2.9 | 212 | 13 | US-09-540-208-31268 | Sequence 31268, A | 273 | 17 | 2.9 | 250 | 21 | US-09-540-766-16329 | Sequence 36329, A |
| 201 | 17 | 2.9 | 212 | 34 | US-60-023-339-1972 | Sequence 1972, Ap | 274 | 17 | 2.9 | 250 | 35 | US-60-036-145-731 | Sequence 731, App |
| 202 | 17 | 2.9 | 215 | 21 | US-09-540-764-3722 | Sequence 3722, Ap | 275 | 17 | 2.9 | 251 | 14 | US-09-035-171-21 | Sequence 21, App |
| 203 | 17 | 2.9 | 216 | 17 | US-09-394-745-53715 | Sequence 53715, A | 276 | 17 | 2.9 | 251 | 17 | US-09-304-517A-136266 | Sequence 136266 |
| 204 | 17 | 2.9 | 218 | 16 | US-09-294-086A-5439 | Sequence 5439, Ap | 277 | 17 | 2.9 | 251 | 21 | US-09-371-146A-136266 | Sequence 136266 |
| 205 | 17 | 2.9 | 218 | 40 | US-60-084-492-5364 | Sequence 5364, Ap | 278 | 17 | 2.9 | 251 | 21 | US-09-540-212A-11572 | Sequence 11572, A |
| 206 | 17 | 2.9 | 222 | 14 | US-09-021-702-8 | Sequence 8, App1 | 279 | 17 | 2.9 | 251 | 35 | US-60-039-416-21 | Sequence 21, App1 |
| 207 | 17 | 2.9 | 222 | 21 | US-09-540-766-6330 | Sequence 6330, Ap | 280 | 17 | 2.9 | 252 | 17 | US-09-329-021-5640 | Sequence 5640, Ap |
| 208 | 17 | 2.9 | 222 | 36 | US-60-041-220-8 | Sequence 8, App1 | 281 | 17 | 2.9 | 252 | 17 | US-60-089-805-5640 | Sequence 5640, Ap |
| 209 | 17 | 2.9 | 223 | 13 | US-08-922-314-1076 | Sequence 1076, Ap | 282 | 17 | 2.9 | 253 | 14 | US-09-022-355-1796 | Sequence 1796, Ap |
| 210 | 17 | 2.9 | 223 | 13 | US-08-922-314A-1076 | Sequence 1076, Ap | 283 | 17 | 2.9 | 253 | 14 | US-09-022-355-1803 | Sequence 1803, Ap |
| 211 | 17 | 2.9 | 223 | 22 | US-09-565-306-78209 | Sequence 78209, A | 284 | 17 | 2.9 | 253 | 14 | US-09-036-589-1307 | Sequence 1307, Ap |
| 212 | 17 | 2.9 | 224 | 13 | US-08-903-469-3395 | Sequence 3395, Ap | 285 | 17 | 2.9 | 253 | 20 | US-09-539-806-14216 | Sequence 14216, A |
| 213 | 17 | 2.9 | 224 | 14 | US-09-015-080-1958 | Sequence 1958, Ap | 286 | 17 | 2.9 | 253 | 21 | US-09-540-208-16509 | Sequence 16509, A |
| 214 | 17 | 2.9 | 224 | 21 | US-09-540-208-41552 | Sequence 41542, A | 287 | 17 | 2.9 | 253 | 21 | US-09-540-212A-7313 | Sequence 7313, Ap |
| 215 | 17 | 2.9 | 224 | 21 | US-09-540-212A-40039 | Sequence 40039, A | 288 | 17 | 2.9 | 253 | 35 | US-60-039-128-1307 | Sequence 1307, Ap |
| 216 | 17 | 2.9 | 224 | 34 | US-60-023-339-1025 | Sequence 1025, Ap | 289 | 17 | 2.9 | 253 | 36 | US-60-044-847-1796 | Sequence 1796, Ap |
| 217 | 17 | 2.9 | 224 | 34 | US-60-025-143-1472 | Sequence 1472, Ap | 290 | 17 | 2.9 | 253 | 36 | US-60-044-847-1803 | Sequence 1803, Ap |
| 218 | 17 | 2.9 | 226 | 13 | US-08-922-314-1188 | Sequence 1188, Ap | 291 | 17 | 2.9 | 255 | 14 | US-09-015-080-1921 | Sequence 1921, Ap |
| 219 | 17 | 2.9 | 226 | 13 | US-08-922-314A-1188 | Sequence 1188, Ap | 292 | 17 | 2.9 | 255 | 16 | US-09-252-691-4939 | Sequence 4939, Ap |
| 220 | 17 | 2.9 | 228 | 13 | US-08-922-315-3101 | Sequence 3101, Ap | 293 | 17 | 2.9 | 255 | 16 | US-09-252-691C-4939 | Sequence 4939, Ap |
| 221 | 17 | 2.9 | 230 | 21 | US-09-540-499-23732 | Sequence 23732, A | 294 | 17 | 2.9 | 255 | 21 | US-09-540-212A-20606 | Sequence 20606, A |
| 222 | 17 | 2.9 | 230 | 42 | US-60-108-396-2048 | Sequence 2048, Ap | 295 | 17 | 2.9 | 256 | 19 | US-09-522-304-5687 | Sequence 5687, Ap |
| 223 | 17 | 2.9 | 232 | 21 | US-08-731-034-390 | Sequence 390, App | 296 | 17 | 2.9 | 257 | 13 | US-08-435-761-1431 | Sequence 1431, Ap |
| 224 | 17 | 2.9 | 232 | 21 | US-09-540-733-2131 | Sequence 2131, Ap | 297 | 17 | 2.9 | 257 | 13 | US-08-903-469-1121 | Sequence 1121, Ap |
| 225 | 17 | 2.9 | 232 | 32 | US-60-004-674-390 | Sequence 390, App | 298 | 17 | 2.9 | 257 | 23 | US-09-540-208-1013 | Sequence 1013, A |
| 226 | 17 | 2.9 | 233 | 15 | US-09-187-860-819 | Sequence 819, App | 299 | 17 | 2.9 | 257 | 23 | US-09-619-643-29916 | Sequence 29916, A |
| 227 | 17 | 2.9 | 236 | 17 | US-09-350-960A-2212 | Sequence 2212, Ap | 300 | 17 | 2.9 | 257 | 25 | US-09-654-617-350120 | Sequence 350120, A |
| 228 | 17 | 2.9 | 236 | 41 | US-09-540-229-158438 | Sequence 158438, A | 301 | 17 | 2.9 | 257 | 27 | US-09-684-016-350120 | Sequence 350120, A |
| 229 | 17 | 2.9 | 236 | 21 | US-60-099-523-2212 | Sequence 2212, Ap | 302 | 17 | 2.9 | 257 | 46 | US-60-145-116-3690 | Sequence 3690, Ap |
| 230 | 17 | 2.9 | 238 | 11 | US-08-731-034-2116 | Sequence 2116, Ap | 303 | 17 | 2.9 | 258 | 12 | US-08-865-291-1612 | Sequence 1612, Ap |
| 231 | 17 | 2.9 | 238 | 21 | US-09-540-733-1091 | Sequence 1091, Ap | 304 | 17 | 2.9 | 258 | 13 | US-08-992-868-2604 | Sequence 2604, Ap |
| 232 | 17 | 2.9 | 238 | 22 | US-09-594-596-2285 | Sequence 2285, Ap | 305 | 17 | 2.9 | 258 | 14 | US-09-056-942-1059 | Sequence 1059, Ap |
| 233 | 17 | 2.9 | 238 | 25 | US-09-654-617-448888 | Sequence 448888, A | 306 | 17 | 2.9 | 258 | 15 | US-09-100-454-1399 | Sequence 1399, Ap |
| 234 | 17 | 2.9 | 238 | 27 | US-09-684-016-448888 | Sequence 448888, A | 307 | 17 | 2.9 | 258 | 16 | US-09-231-925-74 | Sequence 74, App1 |
| 235 | 17 | 2.9 | 238 | 32 | US-60-004-674-2116 | Sequence 2116, Ap | 308 | 17 | 2.9 | 258 | 21 | US-09-539-801-274 | Sequence 274, App |
| 236 | 17 | 2.9 | 240 | 11 | US-08-731-034-104 | Sequence 35, App1 | 309 | 17 | 2.9 | 258 | 21 | US-09-540-212A-15521 | Sequence 15521, A |
| 237 | 17 | 2.9 | 240 | 13 | US-08-986-693-768 | Sequence 768, App | 310 | 17 | 2.9 | 258 | 32 | US-09-540-229-155027 | Sequence 155027, A |
| 238 | 17 | 2.9 | 240 | 13 | US-08-986-693A-768 | Sequence 768, App | 311 | 17 | 2.9 | 258 | 36 | US-60-044-882-1059 | Sequence 1059, Ap |
| 239 | 17 | 2.9 | 240 | 21 | US-09-540-229-54519 | Sequence 54519, A | 312 | 17 | 2.9 | 258 | 37 | US-60-051-749-1059 | Sequence 1059, Ap |
| 240 | 17 | 2.9 | 240 | 21 | US-09-540-733-3800 | Sequence 3800, Ap | 313 | 17 | 2.9 | 260 | 31 | US-08-740-370-890 | Sequence 890, App |
| 241 | 17 | 2.9 | 240 | 32 | US-60-004-674-104 | Sequence 104, App | 314 | 17 | 2.9 | 260 | 32 | US-60-007-127-890 | Sequence 890, App |
| 242 | 17 | 2.9 | 241 | 14 | US-09-015-080-2355 | Sequence 2355, Ap | 315 | 17 | 2.9 | 261 | 11 | US-08-731-034-1461 | Sequence 1461, Ap |
| 243 | 17 | 2.9 | 241 | 14 | US-09-035-171-51 | Sequence 51, App1 | 316 | 17 | 2.9 | 261 | 21 | US-09-540-733-1422 | Sequence 1422, Ap |
| 244 | 17 | 2.9 | 241 | 15 | US-09-114-053-559 | Sequence 559, App | 317 | 17 | 2.9 | 261 | 32 | US-60-004-674-1061 | Sequence 1061, Ap |
| 245 | 17 | 2.9 | 241 | 18 | US-09-420-691-3732 | Sequence 3732, App | 318 | 17 | 2.9 | 263 | 43 | US-09-340-766-10834 | Sequence 70834, A |
| 246 | 17 | 2.9 | 241 | 21 | US-09-540-212A-6342 | Sequence 6342, Ap | 319 | 17 | 2.9 | 263 | 43 | US-60-111-749-1395 | Sequence 1395, A |
| 247 | 17 | 2.9 | 241 | 21 | US-09-540-212A-7553 | Sequence 7553, Ap | 320 | 17 | 2.9 | 264 | 5 | US-08-104-507A-6744 | Sequence 6744, Ap |
| 248 | 17 | 2.9 | 242 | 35 | US-60-039-416-51 | Sequence 51, App1 | 321 | 17 | 2.9 | 264 | 5 | US-08-104-507C-6744 | Sequence 6744, Ap |
| 249 | 17 | 2.9 | 242 | 13 | US-08-901-904-5784 | Sequence 5784, Ap | 322 | 17 | 2.9 | 264 | 12 | US-08-847-659-1569 | Sequence 1569, Ap |
| 250 | 17 | 2.9 | 242 | 21 | US-09-540-208-6288 | Sequence 6288, Ap | 323 | 17 | 2.9 | 264 | 12 | US-08-858-221-2 | Sequence 2, App1 |
| 251 | 17 | 2.9 | 242 | 21 | US-08-731-034-1013 | Sequence 1013, Ap | 324 | 17 | 2.9 | 264 | 14 | US-09-032-271-266 | Sequence 266, App |
| 252 | 17 | 2.9 | 244 | 11 | US-09-540-733-2542 | Sequence 2542, Ap | 325 | 17 | 2.9 | 264 | 31 | US-09-540-229-74017 | Sequence 74017, A |
| 253 | 17 | 2.9 | 244 | 32 | US-60-004-674-1013 | Sequence 1013, Ap | 326 | 17 | 2.9 | 264 | 35 | US-60-038-585-266 | Sequence 266, App |
| 254 | 17 | 2.9 | 244 | 32 | US-08-722-922-29 | Sequence 29, App1 | 327 | 17 | 2.9 | 267 | 21 | US-09-540-212A-39486 | Sequence 39486, A |
| 255 | 17 | 2.9 | 245 | 11 | US-08-731-034-614 | Sequence 614, App | 328 | 17 | 2.9 | 267 | 50 | US-60-184-777-1059 | Sequence 1059, Ap |
| 256 | 17 | 2.9 | 245 | 18 | US-09-478-821-4957 | Sequence 4957, Ap | 329 | 17 | 2.9 | 269 | 22 | US-09-565-109A-2455 | Sequence 2455, Ap |
| 257 | 17 | 2.9 | 245 | 18 | US-09-540-229-154330 | Sequence 154330, Ap | 330 | 17 | 2.9 | 272 | 14 | US-09-066-970-532 | Sequence 532, App |
| 258 | 17 | 2.9 | 245 | 21 | US-09-540-733-2287 | Sequence 2287, Ap | 331 | 17 | 2.9 | 272 | 36 | US-60-044-798-532 | Sequence 532, App |
| 259 | 17 | 2.9 | 245 | 31 | US-60-004-674-1670 | Sequence 1670, Ap | 332 | 17 | 2.9 | 275 | 13 | US-08-731-034-11257 | Sequence 11257, Ap |
| 260 | 17 | 2.9 | 245 | 31 | US-60-004-674-1670 | Sequence 1670, Ap | 333 | 17 | 2.9 | 275 | 13 | US-08-922-315-1164 | Sequence 1164, Ap |
| 261 | 17 | 2.9 | 245 | 43 | US-08-731-034-614 | Sequence 614, App | 334 | 17 | 2.9 | 275 | 15 | US-09-187-860-1192 | Sequence 1192, Ap |
| 262 | 17 | 2.9 | 246 | 14 | US-09-035-171-48 | Sequence 48, App1 | 335 | 17 | 2.9 | 275 | 21 | US-09-540-733-720 | Sequence 720, App |
| 263 | 17 | 2.9 | 246 | 14 | US-09-540-212A-9617 | Sequence 9617, Ap | 336 | 17 | 2.9 | 275 | 32 | US-60-004-674-1257 | Sequence 1257, App |
| 264 | 17 | 2.9 | 246 | 21 | US-09-540-733-2706 | Sequence 2706, Ap | 337 | 17 | 2.9 | 276 | 15 | US-09-129-789-1389 | Sequence 1389, Ap |
| 265 | 17 | 2.9 | 246 | 21 | US-09-540-733-2706 | Sequence 2706, Ap | 338 | 17 | 2.9 | 276 | 21 | US-09-540-229-167205 | Sequence 167205, A |

| | | | | | | | | | | | | | |
|-----|----|-----|-----|----|----------------------|---------------------|-------|----|-----|-----|----|-----------------------|--------------------|
| 339 | 17 | 2.9 | 276 | 21 | US-09-540-766-77676 | Sequence 77676, A | C 412 | 17 | 2.9 | 322 | 16 | US-09-266-640B-3604 | Sequence 3604, Ap |
| 340 | 17 | 2.9 | 276 | 46 | US-60-141-230-731 | Sequence 731, App | 413 | 17 | 2.9 | 322 | 20 | US-09-539-806-8047 | Sequence 8047, Ap |
| 341 | 17 | 2.9 | 277 | 20 | US-09-539-806-14607 | Sequence 14607, A | 414 | 17 | 2.9 | 336 | 17 | US-09-321-214-1436 | Sequence 4636, Ap |
| 342 | 17 | 2.9 | 277 | 21 | US-09-540-766-59323 | Sequence 59323, A | 415 | 17 | 2.9 | 336 | 17 | US-09-359-067-5584 | Sequence 5584, Ap |
| 343 | 17 | 2.9 | 277 | 15 | US-09-167-461-124 | Sequence 124, App | 416 | 17 | 2.9 | 336 | 19 | US-09-516-335-4636 | Sequence 4636, Ap |
| 344 | 17 | 2.9 | 278 | 21 | US-09-540-766-67184 | Sequence 67184, A | 417 | 17 | 2.9 | 336 | 29 | US-09-733-811-4636 | Sequence 4636, Ap |
| 345 | 17 | 2.9 | 280 | 11 | US-08-731-034-1435 | Sequence 1435, Ap | 418 | 17 | 2.9 | 336 | 29 | US-09-733-811A-4636 | Sequence 4636, Ap |
| 346 | 17 | 2.9 | 280 | 21 | US-09-540-733-2847 | Sequence 2847, Ap | 419 | 17 | 2.9 | 337 | 53 | US-60-211-750-30344 | Sequence 30344, A |
| 347 | 17 | 2.9 | 280 | 32 | US-60-004-674-1435 | Sequence 1435, Ap | 420 | 17 | 2.9 | 337 | 16 | US-09-271-490-33557 | Sequence 33557, Ap |
| 348 | 17 | 2.9 | 281 | 12 | US-60-823-505-339 | Sequence 339, App | 421 | 17 | 2.9 | 344 | 50 | US-60-180-489-898 | Sequence 898, App |
| 349 | 17 | 2.9 | 281 | 14 | US-09-036-310-6450 | Sequence 6450, Ap | 422 | 17 | 2.9 | 348 | 19 | US-09-521-640-94962 | Sequence 94962, A |
| 350 | 17 | 2.9 | 281 | 16 | US-09-294-086A-944 | Sequence 944, App | 423 | 17 | 2.9 | 358 | 24 | US-09-637-086A-39732 | Sequence 39732, A |
| 351 | 17 | 2.9 | 281 | 21 | US-09-540-229-19904 | Sequence 19904, A | 424 | 17 | 2.9 | 365 | 51 | US-60-192-733-1665 | Sequence 1665, Ap |
| 352 | 17 | 2.9 | 281 | 21 | US-09-540-229-26772 | Sequence 26772, A | 425 | 17 | 2.9 | 365 | 51 | US-60-194-243-1220 | Sequence 1220, Ap |
| 353 | 17 | 2.9 | 281 | 40 | US-60-084-492-783 | Sequence 783, App | 426 | 17 | 2.9 | 365 | 17 | US-09-362-510-14653 | Sequence 14653, A |
| 354 | 17 | 2.9 | 282 | 25 | US-09-654-617-358073 | Sequence 358073, A | 427 | 17 | 2.9 | 366 | 17 | US-09-362-510-14653 | Sequence 14653, A |
| 355 | 17 | 2.9 | 282 | 27 | US-09-684-016-358073 | Sequence 358073, A | 428 | 17 | 2.9 | 391 | 31 | US-09-872-850-160 | Sequence 160, App |
| 356 | 17 | 2.9 | 284 | 11 | US-08-779-625-56 | Sequence 56, App1 | 429 | 17 | 2.9 | 392 | 17 | US-09-362-510-59818 | Sequence 59818, A |
| 357 | 17 | 2.9 | 284 | 17 | US-09-323-237-111 | Sequence 111, App | 430 | 17 | 2.9 | 392 | 17 | US-09-362-510A-59818 | Sequence 59818, A |
| 358 | 17 | 2.9 | 284 | 22 | US-09-577-410-4954 | Sequence 4954, Ap | 431 | 17 | 2.9 | 393 | 24 | US-09-620-393-1212 | Sequence 1212, Ap |
| 359 | 17 | 2.9 | 284 | 33 | US-60-010-158-56 | Sequence 56, App1 | 432 | 17 | 2.9 | 394 | 51 | US-60-197-872-49698 | Sequence 49698, A |
| 360 | 17 | 2.9 | 284 | 40 | US-60-089-804-111 | Sequence 111, App1 | 433 | 17 | 2.9 | 395 | 17 | US-09-306-609-2729 | Sequence 2729, Ap |
| 361 | 17 | 2.9 | 285 | 15 | US-09-149-716B-42 | Sequence 42, App1 | 434 | 17 | 2.9 | 395 | 19 | US-09-522-251-2729 | Sequence 2729, Ap |
| 362 | 17 | 2.9 | 285 | 21 | US-09-540-229-142795 | Sequence 142795, A | 435 | 17 | 2.9 | 397 | 16 | US-09-270-849B-170029 | Sequence 170029, A |
| 363 | 17 | 2.9 | 285 | 21 | US-09-540-499-22079 | Sequence 22079, A | 436 | 17 | 2.9 | 398 | 50 | US-60-189-657-15584 | Sequence 15584, A |
| 364 | 17 | 2.9 | 286 | 21 | US-09-540-208-69920 | Sequence 69920, A | 437 | 17 | 2.9 | 400 | 17 | US-09-321-214-13011 | Sequence 13011, A |
| 365 | 17 | 2.9 | 287 | 21 | US-09-540-499-21348 | Sequence 21348, A | 438 | 17 | 2.9 | 400 | 19 | US-09-516-335-13011 | Sequence 13011, A |
| 366 | 17 | 2.9 | 287 | 24 | US-09-637-886-2609 | Sequence 2609, App | 439 | 17 | 2.9 | 400 | 29 | US-09-733-811-13011 | Sequence 13011, A |
| 367 | 17 | 2.9 | 289 | 14 | US-09-056-942-310 | Sequence 310, App | 440 | 17 | 2.9 | 400 | 29 | US-09-733-811A-13011 | Sequence 13011, A |
| 368 | 17 | 2.9 | 289 | 21 | US-09-540-212A-40653 | Sequence 40653, A | 441 | 17 | 2.9 | 402 | 16 | US-09-348-797-15681 | Sequence 15681, A |
| 369 | 17 | 2.9 | 289 | 36 | US-60-044-082-310 | Sequence 310, App | 442 | 17 | 2.9 | 402 | 16 | US-09-348-956-4316 | Sequence 4316, Ap |
| 370 | 17 | 2.9 | 290 | 15 | US-09-167-461-69 | Sequence 69, App1 | 443 | 17 | 2.9 | 406 | 22 | US-09-565-308A-2458 | Sequence 2458, Ap |
| 371 | 17 | 2.9 | 290 | 21 | US-09-540-766-72119 | Sequence 72119, A | 444 | 17 | 2.9 | 411 | 23 | US-09-619-643-8852 | Sequence 8852, Ap |
| 372 | 17 | 2.9 | 290 | 52 | US-60-207-458-8822 | Sequence 8822, Ap | 445 | 17 | 2.9 | 411 | 46 | US-60-145-485-1939 | Sequence 1939, Ap |
| 373 | 17 | 2.9 | 292 | 21 | US-09-540-229-183485 | Sequence 183485, A | 446 | 17 | 2.9 | 414 | 14 | US-09-078-617-2752 | Sequence 2752, Ap |
| 374 | 17 | 2.9 | 293 | 14 | US-09-075-126-1345 | Sequence 1345, Ap | 447 | 17 | 2.9 | 414 | 17 | US-09-359-067-11090 | Sequence 11090, A |
| 375 | 17 | 2.9 | 293 | 21 | US-09-540-208-49528 | Sequence 49528, A | 448 | 17 | 2.9 | 417 | 29 | US-09-326-175-1039 | Sequence 1039, Ap |
| 376 | 17 | 2.9 | 293 | 36 | US-60-048-726-1345 | Sequence 1345, Ap | 449 | 17 | 2.9 | 418 | 21 | US-09-540-235-3086 | Sequence 3086, Ap |
| 377 | 17 | 2.9 | 296 | 14 | US-09-065-511-2418 | Sequence 2418, Ap | 450 | 17 | 2.9 | 418 | 21 | US-60-128-433-4700 | Sequence 4700, Ap |
| 378 | 17 | 2.9 | 296 | 31 | US-09-540-229-186413 | Sequence 186413, A | 451 | 17 | 2.9 | 418 | 58 | US-60-207-458-12481 | Sequence 12481, A |
| 379 | 17 | 2.9 | 296 | 30 | US-09-770-175-4043 | Sequence 4043, Ap | 452 | 17 | 2.9 | 419 | 42 | US-60-162-747-6251 | Sequence 6251, Ap |
| 380 | 17 | 2.9 | 298 | 15 | US-09-128-841-4141 | Sequence 4141, Ap | 453 | 17 | 2.9 | 421 | 19 | US-09-528-409-50541 | Sequence 50541, A |
| 381 | 17 | 2.9 | 298 | 21 | US-09-540-229-169206 | Sequence 169206, A | 454 | 17 | 2.9 | 421 | 53 | US-60-211-750-181842 | Sequence 181842, A |
| 382 | 17 | 2.9 | 298 | 38 | US-60-066-010-2340 | Sequence 2340, Ap | 455 | 17 | 2.9 | 424 | 14 | US-09-076-897-5766 | Sequence 5766, Ap |
| 383 | 17 | 2.9 | 299 | 14 | US-09-083-894-2130 | Sequence 2130, Ap | 456 | 17 | 2.9 | 424 | 14 | US-09-076-897A-5766 | Sequence 5766, Ap |
| 384 | 17 | 2.9 | 300 | 11 | US-08-731-034-1375 | Sequence 1375, Ap | 457 | 17 | 2.9 | 424 | 14 | US-09-076-897B-5766 | Sequence 5766, Ap |
| 385 | 17 | 2.9 | 300 | 19 | US-09-036-591-1707 | Sequence 1707, Ap | 458 | 17 | 2.9 | 424 | 14 | US-09-076-897C-5766 | Sequence 5766, Ap |
| 386 | 17 | 2.9 | 300 | 19 | US-09-500-782-971 | Sequence 971, App | 459 | 17 | 2.9 | 424 | 52 | US-60-208-063-10717 | Sequence 10717, A |
| 387 | 17 | 2.9 | 300 | 21 | US-09-540-733-4263 | Sequence 4263, Ap | 460 | 17 | 2.9 | 425 | 19 | US-09-528-409-107156 | Sequence 107156, A |
| 388 | 17 | 2.9 | 300 | 21 | US-09-540-766-5496 | Sequence 5496, Ap | 461 | 17 | 2.9 | 425 | 25 | US-09-652-911-2079 | Sequence 2079, Ap |
| 389 | 17 | 2.9 | 300 | 32 | US-60-004-674-1375 | Sequence 1375, Ap | 462 | 17 | 2.9 | 425 | 25 | US-09-654-617-269791 | Sequence 269791, A |
| 390 | 17 | 2.9 | 300 | 35 | US-60-036-549-11707 | Sequence 11707, App | 463 | 17 | 2.9 | 425 | 27 | US-09-684-016-269791 | Sequence 269791, A |
| 391 | 17 | 2.9 | 300 | 43 | US-60-119-705-971 | Sequence 971, App | 464 | 17 | 2.9 | 426 | 22 | US-09-572-409-50257 | Sequence 50257, A |
| 392 | 17 | 2.9 | 301 | 14 | US-09-040-266-1869 | Sequence 1869, App | 465 | 17 | 2.9 | 426 | 50 | US-60-189-657-8116 | Sequence 8116, Ap |
| 393 | 17 | 2.9 | 302 | 13 | US-08-936-937-1473 | Sequence 1473, Ap | 466 | 17 | 2.9 | 429 | 49 | US-60-172-367-1136 | Sequence 1136, App |
| 394 | 17 | 2.9 | 302 | 15 | US-09-107-909-1196 | Sequence 1196, Ap | 467 | 17 | 2.9 | 430 | 16 | US-09-234-611-4385 | Sequence 4385, Ap |
| 395 | 17 | 2.9 | 302 | 21 | US-09-540-212A-5492 | Sequence 5492, App | 468 | 17 | 2.9 | 430 | 16 | US-09-248-797-43363 | Sequence 43363, A |
| 396 | 17 | 2.9 | 302 | 32 | US-60-008-131-1473 | Sequence 1473, Ap | 469 | 17 | 2.9 | 430 | 17 | US-09-394-745-52273 | Sequence 52273, A |
| 397 | 17 | 2.9 | 303 | 14 | US-09-070-694-151 | Sequence 151, App | 470 | 17 | 2.9 | 432 | 22 | US-09-565-306-83030 | Sequence 83030, A |
| 398 | 17 | 2.9 | 303 | 36 | US-60-044-030-151 | Sequence 249, App | 471 | 17 | 2.9 | 432 | 22 | US-09-565-306-83030 | Sequence 83030, A |
| 399 | 17 | 2.9 | 307 | 19 | US-09-516-278-249 | Sequence 249, App | 472 | 17 | 2.9 | 445 | 19 | US-09-515-128-21541 | Sequence 21541, A |
| 400 | 17 | 2.9 | 307 | 29 | US-09-736-119-680 | Sequence 680, App | 473 | 17 | 2.9 | 445 | 22 | US-09-572-409-66461 | Sequence 66461, A |
| 401 | 17 | 2.9 | 309 | 14 | US-09-042-629-317 | Sequence 317, App | 474 | 17 | 2.9 | 450 | 17 | US-09-371-805-5152 | Sequence 5152, Ap |
| 402 | 17 | 2.9 | 309 | 20 | US-09-539-806-10103 | Sequence 10103, A | 475 | 17 | 2.9 | 450 | 17 | US-09-371-805-5152 | Sequence 5152, Ap |
| 403 | 17 | 2.9 | 310 | 12 | US-08-808-443A-28115 | Sequence 28115, Ap | 476 | 17 | 2.9 | 450 | 52 | US-60-207-458-93904 | Sequence 93904, A |
| 404 | 17 | 2.9 | 310 | 12 | US-08-880-314-1239 | Sequence 1239, Ap | 477 | 17 | 2.9 | 451 | 19 | US-09-528-409-107157 | Sequence 107157, A |
| 405 | 17 | 2.9 | 310 | 21 | US-09-540-208-25550 | Sequence 25550, A | 478 | 17 | 2.9 | 452 | 22 | US-09-572-409-74513 | Sequence 74513, A |
| 406 | 17 | 2.9 | 310 | 33 | US-60-012-233-2815 | Sequence 2815, Ap | 479 | 17 | 2.9 | 452 | 22 | US-09-572-409-74513 | Sequence 74513, A |
| 407 | 17 | 2.9 | 317 | 17 | US-09-323-239-2015 | Sequence 2015, Ap | 480 | 17 | 2.9 | 456 | 25 | US-09-654-617-161371 | Sequence 161371, A |
| 408 | 17 | 2.9 | 317 | 17 | US-09-323-239-2015 | Sequence 2015, Ap | 481 | 17 | 2.9 | 456 | 27 | US-09-684-409-42982 | Sequence 42982, A |
| 409 | 17 | 2.9 | 322 | 13 | US-08-839-968-1234 | Sequence 1234, Ap | 482 | 17 | 2.9 | 457 | 16 | US-09-528-409-42982 | Sequence 42982, A |
| 410 | 17 | 2.9 | 322 | 13 | US-08-951-198-1234 | Sequence 1234, Ap | 483 | 17 | 2.9 | 467 | 53 | US-60-217-080-19231 | Sequence 19231, A |
| 411 | 17 | 2.9 | 322 | 16 | US-09-266-640A-3604 | Sequence 3604, Ap | 484 | 17 | 2.9 | 467 | 16 | US-09-289-766-31425 | Sequence 31425, A |

| | | | | | | | | | | | | | |
|-----|----|-----|-----|----|----------------------|--------------------|-------|----|-----|------|----|-----------------------|--------------------|
| 485 | 17 | 2.9 | 467 | 17 | US-09-332-782-26358 | Sequence 26358, A | c 558 | 17 | 2.9 | 691 | 28 | US-09-702-134-2532 | Sequence 2532, Ap |
| 486 | 17 | 2.9 | 467 | 21 | US-09-540-229-129339 | Sequence 129339, A | 559 | 17 | 2.9 | 706 | 27 | US-09-699-998-4636 | Sequence 4636, Ap |
| 487 | 17 | 2.9 | 467 | 29 | US-09-737-223-2638 | Sequence 2638, A | 560 | 17 | 2.9 | 724 | 49 | US-09-172-362-3158 | Sequence 3158, Ap |
| 488 | 17 | 2.9 | 474 | 17 | US-09-306-609-9398 | Sequence 9398, Ap | 561 | 17 | 2.9 | 731 | 1 | PCT-US01-10485A-72 | Sequence 72, Appl |
| 489 | 17 | 2.9 | 474 | 19 | US-09-522-251-9398 | Sequence 9398, Ap | 562 | 17 | 2.9 | 731 | 1 | PCT-US01-10485A-72 | Sequence 72, Appl |
| 490 | 17 | 2.9 | 477 | 16 | US-09-270-849B-96172 | Sequence 96172, A | 563 | 17 | 2.9 | 765 | 22 | US-09-565-309A-53882 | Sequence 53882, A |
| 491 | 17 | 2.9 | 490 | 30 | US-09-770-175-4202 | Sequence 4202, Ap | 564 | 17 | 2.9 | 784 | 16 | US-09-270-849B-160732 | Sequence 160732, A |
| 492 | 17 | 2.9 | 493 | 25 | US-09-654-617-198292 | Sequence 198292, A | 565 | 17 | 2.9 | 786 | 18 | US-09-489-039A-1447 | Sequence 1447, Ap |
| 493 | 17 | 2.9 | 493 | 27 | US-09-684-016-198292 | Sequence 198292, A | 566 | 17 | 2.9 | 791 | 16 | US-09-270-849B-115904 | Sequence 115904, A |
| 494 | 17 | 2.9 | 494 | 52 | US-09-684-016-198292 | Sequence 198292, A | 567 | 17 | 2.9 | 800 | 1 | PCT-US00-12788-32 | Sequence 32, Appl |
| 495 | 17 | 2.9 | 496 | 25 | US-09-644-871-485 | Sequence 485, App | 568 | 17 | 2.9 | 806 | 30 | US-09-770-175-3164 | Sequence 3164, Ap |
| 496 | 17 | 2.9 | 496 | 25 | US-09-644-871-485 | Sequence 485, App | 569 | 17 | 2.9 | 906 | 30 | US-09-770-175-8627 | Sequence 8627, Ap |
| 497 | 17 | 2.9 | 500 | 25 | US-09-654-617-457275 | Sequence 457275, A | 570 | 17 | 2.9 | 957 | 17 | US-09-371-168-7394 | Sequence 7394, Ap |
| 498 | 17 | 2.9 | 500 | 27 | US-09-684-016-457275 | Sequence 457275, A | 571 | 17 | 2.9 | 957 | 25 | US-09-644-870-9050 | Sequence 9050, Ap |
| 499 | 17 | 2.9 | 500 | 52 | US-09-684-016-457275 | Sequence 457275, A | 572 | 17 | 2.9 | 957 | 25 | US-09-644-870-9050 | Sequence 9050, Ap |
| 500 | 17 | 2.9 | 501 | 51 | US-09-684-016-457275 | Sequence 457275, A | 573 | 17 | 2.9 | 957 | 31 | US-09-644-873-9481 | Sequence 9481, Ap |
| 501 | 17 | 2.9 | 513 | 23 | US-09-605-699-16043 | Sequence 16043, A | 574 | 17 | 2.9 | 957 | 31 | US-09-801-833-7394 | Sequence 7394, Ap |
| 502 | 17 | 2.9 | 513 | 23 | US-09-605-699-16043 | Sequence 16043, A | 575 | 17 | 2.9 | 997 | 16 | US-09-270-849B-101870 | Sequence 101870, A |
| 503 | 17 | 2.9 | 521 | 21 | US-09-540-229-117126 | Sequence 117126, A | 576 | 17 | 2.9 | 1015 | 50 | US-09-682-182-616-483 | Sequence 101870, A |
| 504 | 17 | 2.9 | 522 | 25 | US-09-652-911-8081 | Sequence 8081, Ap | 577 | 17 | 2.9 | 1152 | 16 | US-09-252-991A-13050 | Sequence 13050, A |
| 505 | 17 | 2.9 | 527 | 22 | US-09-565-309A-6558 | Sequence 6558, Ap | 578 | 17 | 2.9 | 1170 | 16 | US-09-252-991A-3174 | Sequence 3174, Ap |
| 506 | 17 | 2.9 | 527 | 22 | US-09-565-309A-6558 | Sequence 6558, Ap | 579 | 17 | 2.9 | 1224 | 48 | US-09-252-991A-3174 | Sequence 3174, Ap |
| 507 | 17 | 2.9 | 530 | 25 | US-09-654-617-268058 | Sequence 268058, A | 580 | 17 | 2.9 | 1233 | 16 | US-09-164-285-6287 | Sequence 6287, Ap |
| 508 | 17 | 2.9 | 536 | 17 | US-09-684-016-268058 | Sequence 268058, A | 581 | 17 | 2.9 | 1233 | 16 | US-09-252-991A-13225 | Sequence 13225, A |
| 509 | 17 | 2.9 | 536 | 17 | US-09-306-609-9785 | Sequence 9785, Ap | 582 | 17 | 2.9 | 1278 | 22 | US-09-565-309A-57226 | Sequence 57226, A |
| 510 | 17 | 2.9 | 545 | 45 | US-09-522-251-9785 | Sequence 9785, Ap | 583 | 17 | 2.9 | 1302 | 16 | US-09-252-991A-5267 | Sequence 5267, Ap |
| 511 | 17 | 2.9 | 549 | 27 | US-09-699-997-6689 | Sequence 6689, Ap | 584 | 17 | 2.9 | 1338 | 1 | PCT-US99-03265-153 | Sequence 153, App |
| 512 | 17 | 2.9 | 554 | 46 | US-09-699-997-6689 | Sequence 6689, Ap | 585 | 17 | 2.9 | 1338 | 13 | US-09-942-942-153 | Sequence 153, App |
| 513 | 17 | 2.9 | 557 | 22 | US-09-684-016-268058 | Sequence 268058, A | 586 | 17 | 2.9 | 1338 | 13 | US-09-942-942-153 | Sequence 153, App |
| 514 | 17 | 2.9 | 560 | 52 | US-09-655-309A-44765 | Sequence 44765, A | 587 | 17 | 2.9 | 1338 | 14 | US-09-942-942-153 | Sequence 153, App |
| 515 | 17 | 2.9 | 562 | 52 | US-09-655-309A-44765 | Sequence 44765, A | 588 | 17 | 2.9 | 1338 | 14 | US-09-942-942-153 | Sequence 153, App |
| 516 | 17 | 2.9 | 567 | 17 | US-09-830-44842 | Sequence 44842, A | 589 | 17 | 2.9 | 1338 | 14 | US-09-025-197-158 | Sequence 158, App |
| 517 | 17 | 2.9 | 567 | 17 | US-09-371-435-110 | Sequence 110, App | 590 | 17 | 2.9 | 1338 | 14 | US-09-025-197-158 | Sequence 158, App |
| 518 | 17 | 2.9 | 567 | 30 | US-09-371-435-110 | Sequence 110, App | 591 | 17 | 2.9 | 1338 | 14 | US-09-025-197-158 | Sequence 158, App |
| 519 | 17 | 2.9 | 570 | 50 | US-09-783-062-110 | Sequence 110, App | 592 | 17 | 2.9 | 1338 | 14 | US-09-025-197-158 | Sequence 158, App |
| 520 | 17 | 2.9 | 572 | 25 | US-09-180-489-899 | Sequence 899, App | 593 | 17 | 2.9 | 1338 | 14 | US-09-025-197-158 | Sequence 158, App |
| 521 | 17 | 2.9 | 572 | 25 | US-09-654-617-129789 | Sequence 129789, A | 594 | 17 | 2.9 | 1388 | 48 | US-09-703-708-9484 | Sequence 9484, Ap |
| 522 | 17 | 2.9 | 572 | 25 | US-09-654-617-129789 | Sequence 129789, A | 595 | 17 | 2.9 | 1388 | 48 | US-09-703-708-9484 | Sequence 9484, Ap |
| 523 | 17 | 2.9 | 572 | 25 | US-09-654-617-129789 | Sequence 129789, A | 596 | 17 | 2.9 | 1409 | 26 | US-09-164-320-9484 | Sequence 9484, Ap |
| 524 | 17 | 2.9 | 579 | 22 | US-09-649-164-4158 | Sequence 129329, A | 597 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 525 | 17 | 2.9 | 586 | 42 | US-09-649-164-4158 | Sequence 129329, A | 598 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 526 | 17 | 2.9 | 588 | 16 | US-09-565-309A-6559 | Sequence 6559, Ap | 599 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 527 | 17 | 2.9 | 588 | 16 | US-09-565-309A-6559 | Sequence 6559, Ap | 600 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 528 | 17 | 2.9 | 590 | 49 | US-09-171-481-699 | Sequence 659, App | 601 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 529 | 17 | 2.9 | 606 | 49 | US-09-252-991A-5208 | Sequence 5208, Ap | 602 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 530 | 17 | 2.9 | 606 | 49 | US-09-252-991A-5208 | Sequence 5208, Ap | 603 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 531 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 604 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 532 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 605 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 533 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 606 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 534 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 607 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 535 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 608 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 536 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 609 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 537 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 610 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 538 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 611 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 539 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 612 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 540 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 613 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 541 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 614 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 542 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 615 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 543 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 616 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 544 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 617 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 545 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 618 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 546 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 619 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 547 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 620 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 548 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 621 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 549 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 622 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 550 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 623 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 551 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 624 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 552 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 625 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 553 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 626 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 554 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 627 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 555 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 628 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 556 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 629 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |
| 557 | 17 | 2.9 | 618 | 22 | US-09-173-462-33 | Sequence 854, App | 630 | 17 | 2.9 | 1461 | 27 | US-09-668-683-9423 | Sequence 9423, Ap |

| | | | | | | | | | | | | | |
|-------|----|-----|--------|----|----------------------|-------------------|-------|----|-----|---------|----|-----------------------|--------------------|
| C 631 | 17 | 2.9 | 4571 | 25 | US-09-644-867-7266 | Sequence 7266, Ap | 704 | 17 | 2.9 | 165951 | 19 | US-09-528-237A-1761 | Sequence 1761, Ap |
| C 632 | 17 | 2.9 | 4571 | 25 | US-09-649-166-5716 | Sequence 5716, Ap | 705 | 17 | 2.9 | 167824 | 53 | US-60-212-664-215 | Sequence 215, App |
| C 633 | 17 | 2.9 | 4571 | 25 | US-09-652-914-9395 | Sequence 9395, Ap | 706 | 17 | 2.9 | 170179 | 53 | US-60-212-664-120 | Sequence 120, App |
| C 634 | 17 | 2.9 | 4571 | 27 | US-09-698-010-13248 | Sequence 13248, A | C 707 | 17 | 2.9 | 190628 | 53 | US-60-216-770-25 | Sequence 25, App1 |
| C 635 | 17 | 2.9 | 4571 | 27 | US-09-698-012-8079 | Sequence 8079, Ap | C 708 | 17 | 2.9 | 191626 | 53 | US-60-212-664-111 | Sequence 111, App1 |
| C 636 | 17 | 2.9 | 4571 | 28 | US-09-710-286-3150 | Sequence 3150, Ap | C 709 | 17 | 2.9 | 200624 | 56 | US-60-248-592-14 | Sequence 14, App1 |
| C 637 | 17 | 2.9 | 4571 | 29 | US-09-726-788-6338 | Sequence 6338, Ap | C 710 | 17 | 2.9 | 200676 | 56 | US-60-245-228-162 | Sequence 162, App1 |
| C 638 | 17 | 2.9 | 4571 | 29 | US-09-726-802-2206 | Sequence 2206, Ap | C 711 | 17 | 2.9 | 200676 | 57 | US-60-258-272-65 | Sequence 65, App1 |
| C 639 | 17 | 2.9 | 5276 | 24 | US-09-620-392-54391 | Sequence 54391, A | C 712 | 17 | 2.9 | 4403765 | 15 | US-09-103-840A-2 | Sequence 2, App11 |
| C 640 | 17 | 2.9 | 5603 | 49 | US-60-172-373-4468 | Sequence 4468, Ap | C 713 | 17 | 2.9 | 4411529 | 15 | US-09-103-840A-1 | Sequence 1, App11 |
| C 641 | 17 | 2.9 | 5727 | 24 | US-09-620-392-58663 | Sequence 58663, A | C 714 | 16 | 2.7 | 25 | 55 | US-60-233-166-388596 | Sequence 388596, |
| C 642 | 17 | 2.9 | 6003 | 30 | US-09-770-175-8764 | Sequence 8764, Ap | C 715 | 16 | 2.7 | 25 | 55 | US-60-233-166-395271 | Sequence 395271, |
| C 643 | 17 | 2.9 | 6081 | 24 | US-09-620-392-47347 | Sequence 47347, A | C 716 | 16 | 2.7 | 25 | 55 | US-60-233-620-84989 | Sequence 84989, A |
| C 644 | 17 | 2.9 | 6081 | 28 | US-09-702-134-20091 | Sequence 20091, A | C 717 | 16 | 2.7 | 25 | 55 | US-60-234-017-62378 | Sequence 62378, A |
| C 645 | 17 | 2.9 | 6804 | 24 | US-09-620-392-41879 | Sequence 41879, A | C 718 | 16 | 2.7 | 25 | 55 | US-60-234-017-247278 | Sequence 247278, A |
| C 646 | 17 | 2.9 | 6982 | 28 | US-09-702-134-18646 | Sequence 18646, A | C 719 | 16 | 2.7 | 41 | 18 | US-09-404-520-36397 | Sequence 36397, A |
| C 647 | 17 | 2.9 | 7015 | 24 | US-09-620-392-49807 | Sequence 49807, A | C 720 | 16 | 2.7 | 51 | 49 | US-60-172-373-1937 | Sequence 1937, Ap |
| C 648 | 17 | 2.9 | 7195 | 49 | US-60-173-464-22222 | Sequence 22222, A | C 721 | 16 | 2.7 | 72 | 24 | US-09-628-860-3217 | Sequence 3217, Ap |
| C 649 | 17 | 2.9 | 7195 | 51 | US-60-191-637-35606 | Sequence 35606, A | C 722 | 16 | 2.7 | 136 | 16 | US-09-283-466-1959 | Sequence 1959, Ap |
| C 650 | 17 | 2.9 | 7195 | 51 | US-60-191-681-27929 | Sequence 27929, A | C 723 | 16 | 2.7 | 119 | 16 | US-60-236-359-14114 | Sequence 14114, A |
| C 651 | 17 | 2.9 | 9134 | 56 | US-60-248-830-4-5 | Sequence 5, App11 | C 724 | 16 | 2.7 | 147 | 25 | US-09-654-617-380381 | Sequence 380381, |
| C 652 | 17 | 2.9 | 9723 | 24 | US-09-620-392-50479 | Sequence 50479, A | C 725 | 16 | 2.7 | 147 | 25 | US-09-684-016-380381 | Sequence 380381, |
| C 653 | 17 | 2.9 | 9723 | 28 | US-09-702-134-27209 | Sequence 27209, A | C 726 | 16 | 2.7 | 150 | 22 | US-09-572-409-33672 | Sequence 33672, A |
| C 654 | 17 | 2.9 | 9896 | 24 | US-09-620-392-31492 | Sequence 31492, A | C 727 | 16 | 2.7 | 163 | 25 | US-09-654-617-98770 | Sequence 98770, A |
| C 655 | 17 | 2.9 | 9928 | 24 | US-09-620-392-10823 | Sequence 10823, A | C 728 | 16 | 2.7 | 163 | 27 | US-09-684-016-98770 | Sequence 98770, A |
| C 656 | 17 | 2.9 | 9928 | 28 | US-09-702-134-12531 | Sequence 12531, A | C 729 | 16 | 2.7 | 179 | 7 | US-08-395-244-626 | Sequence 626, App |
| C 657 | 17 | 2.9 | 10265 | 26 | US-09-663-779-1150 | Sequence 1190, Ap | C 730 | 16 | 2.7 | 179 | 13 | US-08-972-899-626 | Sequence 626, App |
| C 658 | 17 | 2.9 | 11212 | 1 | PCT-US00-00913-1 | Sequence 1, App11 | C 731 | 16 | 2.7 | 179 | 20 | US-09-534-845-23701 | Sequence 23701, A |
| C 659 | 17 | 2.9 | 11212 | 16 | US-09-229-862-1 | Sequence 1, App11 | C 732 | 16 | 2.7 | 186 | 16 | US-09-270-849B-172454 | Sequence 172454, |
| C 660 | 17 | 2.9 | 11212 | 18 | US-09-482-788-1 | Sequence 1, App11 | C 733 | 16 | 2.7 | 186 | 23 | US-09-605-070-18448 | Sequence 18448, A |
| C 661 | 17 | 2.9 | 12917 | 28 | US-09-441-340-2 | Sequence 2, App11 | C 734 | 16 | 2.7 | 187 | 53 | US-60-217-080-21303 | Sequence 21303, A |
| C 662 | 17 | 2.9 | 12917 | 28 | US-09-702-134-28365 | Sequence 28365, A | C 735 | 16 | 2.7 | 194 | 53 | US-60-207-458-33168 | Sequence 33168, A |
| C 663 | 17 | 2.9 | 13085 | 24 | US-09-620-392-10510 | Sequence 10510, A | C 736 | 16 | 2.7 | 197 | 17 | US-09-304-517A-8198 | Sequence 8198, Ap |
| C 664 | 17 | 2.9 | 15611 | 18 | US-09-441-340-1 | Sequence 1, App11 | C 737 | 16 | 2.7 | 197 | 17 | US-09-371-146A-8198 | Sequence 8198, Ap |
| C 665 | 17 | 2.9 | 17705 | 24 | US-09-620-392-45583 | Sequence 45583, A | C 738 | 16 | 2.7 | 197 | 25 | US-09-654-617-360140 | Sequence 360140, |
| C 666 | 17 | 2.9 | 18549 | 40 | US-60-082-302-646 | Sequence 646, App | C 739 | 16 | 2.7 | 197 | 25 | US-09-654-617-363016 | Sequence 363016, |
| C 667 | 17 | 2.9 | 19029 | 24 | US-09-620-392-45176 | Sequence 45176, A | C 740 | 16 | 2.7 | 197 | 27 | US-09-684-016-363016 | Sequence 363016, |
| C 668 | 17 | 2.9 | 22730 | 24 | US-09-620-392-42209 | Sequence 42209, A | C 741 | 16 | 2.7 | 197 | 27 | US-09-684-016-363016 | Sequence 363016, |
| C 669 | 17 | 2.9 | 22730 | 24 | US-09-620-392-38958 | Sequence 38958, A | C 742 | 16 | 2.7 | 202 | 8 | US-08-451-242A-216 | Sequence 2166, Ap |
| C 670 | 17 | 2.9 | 24325 | 56 | US-60-242-679-451 | Sequence 451, App | C 743 | 16 | 2.7 | 204 | 20 | US-09-534-843-20541 | Sequence 20541, A |
| C 671 | 17 | 2.9 | 24858 | 28 | US-09-620-392-10811 | Sequence 10811, A | C 744 | 16 | 2.7 | 204 | 44 | US-60-125-817-7188 | Sequence 7188, Ap |
| C 672 | 17 | 2.9 | 24858 | 28 | US-09-702-134-16686 | Sequence 16686, A | C 745 | 16 | 2.7 | 203 | 24 | US-09-637-008-12754 | Sequence 12754, A |
| C 673 | 17 | 2.9 | 27890 | 19 | US-09-528-237A-2031 | Sequence 2031, Ap | C 746 | 16 | 2.7 | 219 | 53 | US-60-217-080-17965 | Sequence 17965, A |
| C 674 | 17 | 2.9 | 27952 | 24 | US-09-620-392-44881 | Sequence 44881, A | C 747 | 16 | 2.7 | 218 | 16 | US-09-298-328A-6513 | Sequence 6513, Ap |
| C 675 | 17 | 2.9 | 28719 | 54 | US-60-226-176-217 | Sequence 217, App | C 748 | 16 | 2.7 | 218 | 17 | US-09-304-517A-25260 | Sequence 25260, A |
| C 676 | 17 | 2.9 | 28719 | 55 | US-60-233-468-217 | Sequence 217, App | C 749 | 16 | 2.7 | 218 | 17 | US-09-371-146A-25260 | Sequence 25260, A |
| C 677 | 17 | 2.9 | 31339 | 56 | US-60-248-592-57 | Sequence 57, App1 | C 750 | 16 | 2.7 | 218 | 40 | US-60-085-147-6513 | Sequence 6513, Ap |
| C 678 | 17 | 2.9 | 32189 | 1 | PCT-US01-01334-8672 | Sequence 8672, Ap | C 751 | 16 | 2.7 | 218 | 40 | US-60-087-762-191 | Sequence 762, Ap |
| C 679 | 17 | 2.9 | 32189 | 1 | PCT-US01-01334-9883 | Sequence 9883, Ap | C 752 | 16 | 2.7 | 218 | 41 | US-60-092-036-918 | Sequence 918, Ap |
| C 680 | 17 | 2.9 | 32189 | 1 | PCT-US01-01334-10121 | Sequence 10121, A | C 753 | 16 | 2.7 | 229 | 22 | US-09-553-094-9948 | Sequence 9948, Ap |
| C 681 | 17 | 2.9 | 32189 | 1 | PCT-US01-01338-3291 | Sequence 3291, Ap | C 754 | 16 | 2.7 | 231 | 11 | US-08-727-699-724 | Sequence 724, App |
| C 682 | 17 | 2.9 | 32340 | 56 | US-60-248-830-4 | Sequence 4, App11 | C 755 | 16 | 2.7 | 232 | 12 | US-08-878-503-1736 | Sequence 1736, Ap |
| C 683 | 17 | 2.9 | 32768 | 52 | US-60-207-419-14 | Sequence 14, App1 | C 756 | 16 | 2.7 | 232 | 21 | US-09-540-229-89054 | Sequence 89054, A |
| C 684 | 17 | 2.9 | 32768 | 54 | US-60-229-514-38 | Sequence 38, App1 | C 757 | 16 | 2.7 | 232 | 21 | US-60-020-017-1736 | Sequence 1736, Ap |
| C 685 | 17 | 2.9 | 32807 | 40 | US-60-082-302-735 | Sequence 735, App | C 758 | 16 | 2.7 | 236 | 16 | US-09-266-640A-4502 | Sequence 4502, Ap |
| C 686 | 17 | 2.9 | 33100 | 40 | US-60-082-302-724 | Sequence 724, App | C 759 | 16 | 2.7 | 236 | 16 | US-09-266-640B-4502 | Sequence 4502, Ap |
| C 687 | 17 | 2.9 | 33494 | 24 | US-09-620-392-1214 | Sequence 1214, Ap | C 760 | 16 | 2.7 | 236 | 17 | US-09-304-517A-20081 | Sequence 20081, A |
| C 688 | 17 | 2.9 | 41790 | 28 | US-09-703-708-755 | Sequence 755, App | C 761 | 16 | 2.7 | 236 | 17 | US-09-371-146A-20081 | Sequence 20081, A |
| C 689 | 17 | 2.9 | 41790 | 48 | US-60-164-320-755 | Sequence 755, App | C 762 | 16 | 2.7 | 236 | 40 | US-60-087-762-191 | Sequence 197, App |
| C 690 | 17 | 2.9 | 41790 | 50 | US-60-183-791-755 | Sequence 755, App | C 763 | 16 | 2.7 | 236 | 41 | US-60-092-036-930 | Sequence 930, App |
| C 691 | 17 | 2.9 | 43140 | 28 | US-09-702-134-6083 | Sequence 6083, Ap | C 764 | 16 | 2.7 | 239 | 9 | US-08-504-732-1175 | Sequence 1175, Ap |
| C 692 | 17 | 2.9 | 45860 | 28 | US-09-702-134-8694 | Sequence 8694, Ap | C 765 | 16 | 2.7 | 242 | 17 | US-09-411-999-2075 | Sequence 2075, Ap |
| C 693 | 17 | 2.9 | 46626 | 20 | US-09-534-859-164 | Sequence 164, App | C 766 | 16 | 2.7 | 241 | 16 | US-09-275-082A-7623 | Sequence 7623, Ap |
| C 694 | 17 | 2.9 | 49209 | 28 | US-09-702-134-941 | Sequence 9461, Ap | C 767 | 16 | 2.7 | 241 | 39 | US-60-079-234-7623 | Sequence 7623, Ap |
| C 695 | 17 | 2.9 | 57690 | 57 | US-60-258-273-54 | Sequence 54, App1 | C 768 | 16 | 2.7 | 241 | 39 | US-60-087-762-193 | Sequence 193, App |
| C 696 | 17 | 2.9 | 63957 | 56 | US-60-248-830-3 | Sequence 3, App11 | C 769 | 16 | 2.7 | 242 | 17 | US-09-304-517A-5554 | Sequence 5554, Ap |
| C 697 | 17 | 2.9 | 78385 | 19 | US-09-528-237A-1583 | Sequence 1583, Ap | C 770 | 16 | 2.7 | 242 | 17 | US-09-371-146A-5554 | Sequence 5554, Ap |
| C 698 | 17 | 2.9 | 88784 | 48 | US-60-161-932-397 | Sequence 397, App | C 771 | 16 | 2.7 | 242 | 17 | US-09-304-517A-40381 | Sequence 40381, A |
| C 699 | 17 | 2.9 | 94029 | 20 | US-09-534-859-147 | Sequence 147, App | C 772 | 16 | 2.7 | 244 | 17 | US-09-371-146A-40381 | Sequence 40381, A |
| C 700 | 17 | 2.9 | 96239 | 28 | US-09-702-134-8772 | Sequence 8772, Ap | C 773 | 16 | 2.7 | 245 | 16 | US-09-244-000A-7232 | Sequence 7232, Ap |
| C 701 | 17 | 2.9 | 110232 | 53 | US-60-212-664-311 | Sequence 311, App | C 774 | 16 | 2.7 | 245 | 16 | US-09-263-191-22755 | Sequence 22755, A |
| C 702 | 17 | 2.9 | 145184 | 19 | US-09-528-237A-1748 | Sequence 1748, Ap | C 775 | 16 | 2.7 | 245 | 16 | US-09-263-191-22755 | Sequence 22755, A |
| C 703 | 17 | 2.9 | 148080 | 53 | US-60-212-664-93 | Sequence 93, App1 | C 776 | 16 | 2.7 | 245 | 17 | US-09-321-209-3251 | Sequence 3251, Ap |

| | | | | | | | | | | | | | |
|-----|----|-----|-----|----|-----------------------|--------------------|-----|----|-----|-----|----|-----------------------|--------------------|
| 777 | 16 | 2.7 | 245 | 18 | US-09-411-999-2281 | Sequence 2281, Ap | 850 | 16 | 2.7 | 279 | 17 | US-09-371-146A-289550 | Sequence 289550, |
| 778 | 16 | 2.7 | 245 | 16 | US-60-086-856-3251 | Sequence 3251, Ap | 851 | 16 | 2.7 | 279 | 39 | US-60-079-234-1403 | Sequence 1403, Ap |
| 779 | 16 | 2.7 | 246 | 26 | US-09-663-779-6398 | Sequence 6398, Ap | 852 | 16 | 2.7 | 279 | 40 | US-60-087-762-187 | Sequence 187, App |
| 780 | 16 | 2.7 | 248 | 16 | US-09-244-000A-91200 | Sequence 91200, A | 853 | 16 | 2.7 | 279 | 41 | US-60-092-036-917 | Sequence 917, App |
| 781 | 16 | 2.7 | 248 | 17 | US-09-304-517A-34234 | Sequence 34234, A | 854 | 16 | 2.7 | 280 | 17 | US-09-304-517A-7615 | Sequence 7615, Ap |
| 782 | 16 | 2.7 | 248 | 17 | US-09-371-146A-34234 | Sequence 34234, A | 855 | 16 | 2.7 | 280 | 17 | US-09-371-146A-7615 | Sequence 7615, Ap |
| 783 | 16 | 2.7 | 249 | 20 | US-09-539-800-1818 | Sequence 1818, Ap | 856 | 16 | 2.7 | 282 | 53 | US-60-217-080-17668 | Sequence 17668, A |
| 784 | 16 | 2.7 | 251 | 16 | US-09-283-466-16073 | Sequence 16073, A | 857 | 16 | 2.7 | 283 | 17 | US-09-306-349-34575 | Sequence 34575, A |
| 785 | 16 | 2.7 | 251 | 17 | US-09-304-517A-37631 | Sequence 37631, A | 858 | 16 | 2.7 | 285 | 5 | US-08-196-363-835 | Sequence 835, App |
| 786 | 16 | 2.7 | 251 | 17 | US-09-371-146A-37631 | Sequence 37631, A | 859 | 16 | 2.7 | 285 | 5 | US-08-196-363A-835 | Sequence 835, App |
| 787 | 16 | 2.7 | 251 | 25 | US-09-654-617-85249 | Sequence 85249, A | 860 | 16 | 2.7 | 285 | 16 | US-09-231-925-1976 | Sequence 1976, Ap |
| 788 | 16 | 2.7 | 251 | 27 | US-09-684-016-85249 | Sequence 85249, A | 861 | 16 | 2.7 | 285 | 16 | US-09-306-349-31944 | Sequence 31944, Ap |
| 789 | 16 | 2.7 | 253 | 12 | US-08-866-372-471 | Sequence 471, App | 862 | 16 | 2.7 | 288 | 17 | US-09-654-617-306034 | Sequence 306034, A |
| 790 | 16 | 2.7 | 253 | 21 | US-09-540-212A-14635 | Sequence 14635, A | 863 | 16 | 2.7 | 289 | 25 | US-09-684-016-306034 | Sequence 306034, A |
| 791 | 16 | 2.7 | 253 | 33 | US-60-018-881-471 | Sequence 471, App | 864 | 16 | 2.7 | 289 | 27 | US-08-823-271-1006 | Sequence 1006, Ap |
| 792 | 16 | 2.7 | 255 | 13 | US-08-999-861-1660 | Sequence 1660, Ap | 865 | 16 | 2.7 | 290 | 12 | US-08-823-271-1006 | Sequence 1006, Ap |
| 793 | 16 | 2.7 | 255 | 16 | US-09-263-191-10531 | Sequence 10531, A | 866 | 16 | 2.7 | 290 | 21 | US-09-540-764-36306 | Sequence 36306, A |
| 794 | 16 | 2.7 | 255 | 53 | US-60-217-080-14307 | Sequence 14307, A | 867 | 16 | 2.7 | 290 | 33 | US-60-015-513-1006 | Sequence 1006, Ap |
| 795 | 16 | 2.7 | 256 | 17 | US-09-306-349-15749 | Sequence 15749, A | 868 | 16 | 2.7 | 293 | 13 | US-08-985-356-229 | Sequence 229, App |
| 796 | 16 | 2.7 | 258 | 22 | US-09-552-086-19696 | Sequence 19696, A | 869 | 16 | 2.7 | 293 | 13 | US-08-985-356A-229 | Sequence 229, App |
| 797 | 16 | 2.7 | 262 | 23 | US-09-605-700-17157 | Sequence 17157, A | 870 | 16 | 2.7 | 295 | 17 | US-09-304-517A-46819 | Sequence 46819, A |
| 798 | 16 | 2.7 | 262 | 18 | US-09-436-711-3514 | Sequence 3514, A | 871 | 16 | 2.7 | 295 | 17 | US-09-371-146A-46819 | Sequence 46819, A |
| 799 | 16 | 2.7 | 263 | 21 | US-09-540-229-106853 | Sequence 106853, A | 872 | 16 | 2.7 | 295 | 18 | US-09-411-999-1984 | Sequence 1984, Ap |
| 800 | 16 | 2.7 | 263 | 40 | US-60-087-762-185 | Sequence 185, App | 873 | 16 | 2.7 | 295 | 52 | US-60-207-458-75596 | Sequence 75596, A |
| 801 | 16 | 2.7 | 263 | 40 | US-60-087-762-189 | Sequence 189, App | 874 | 16 | 2.7 | 296 | 53 | US-60-217-080-17924 | Sequence 17924, A |
| 802 | 16 | 2.7 | 263 | 41 | US-60-092-036-920 | Sequence 920, App | 875 | 16 | 2.7 | 297 | 16 | US-09-215-604A-133 | Sequence 133, App |
| 803 | 16 | 2.7 | 263 | 41 | US-60-092-036-927 | Sequence 927, App | 876 | 16 | 2.7 | 297 | 16 | US-09-215-604A-133 | Sequence 133, App |
| 804 | 16 | 2.7 | 264 | 18 | US-09-436-711-4378 | Sequence 4378, App | 877 | 16 | 2.7 | 298 | 17 | US-09-304-517A-42694 | Sequence 42694, Ap |
| 805 | 16 | 2.7 | 264 | 40 | US-60-087-762-190 | Sequence 190, App | 878 | 16 | 2.7 | 298 | 17 | US-09-371-146A-42694 | Sequence 42694, A |
| 806 | 16 | 2.7 | 264 | 41 | US-60-092-036-924 | Sequence 924, App | 879 | 16 | 2.7 | 298 | 18 | US-09-460-593B-1898 | Sequence 1898, Ap |
| 807 | 16 | 2.7 | 265 | 17 | US-09-112-430-2487 | Sequence 2487, Ap | 880 | 16 | 2.7 | 302 | 5 | US-08-196-482A-1286 | Sequence 1286, Ap |
| 808 | 16 | 2.7 | 265 | 17 | US-09-304-517A-34740 | Sequence 34740, A | 881 | 16 | 2.7 | 302 | 5 | US-08-196-482A-1286 | Sequence 1286, Ap |
| 809 | 16 | 2.7 | 265 | 17 | US-09-371-146A-34740 | Sequence 34740, A | 882 | 16 | 2.7 | 302 | 25 | US-09-654-617-381867 | Sequence 381867, A |
| 810 | 16 | 2.7 | 265 | 18 | US-09-478-821-2487 | Sequence 2487, Ap | 883 | 16 | 2.7 | 302 | 27 | US-09-684-016-381867 | Sequence 381867, A |
| 811 | 16 | 2.7 | 265 | 20 | US-09-534-855-1147 | Sequence 1147, Ap | 884 | 16 | 2.7 | 302 | 50 | US-60-182-467-247 | Sequence 247, App |
| 812 | 16 | 2.7 | 267 | 16 | US-09-244-000A-6476 | Sequence 6476, Ap | 885 | 16 | 2.7 | 302 | 53 | US-60-217-080-13550 | Sequence 13550, A |
| 813 | 16 | 2.7 | 267 | 18 | US-09-436-711-3267 | Sequence 3267, Ap | 886 | 16 | 2.7 | 303 | 14 | US-09-021-095-1406 | Sequence 1406, Ap |
| 814 | 16 | 2.7 | 267 | 40 | US-60-087-762-194 | Sequence 194, App | 887 | 16 | 2.7 | 303 | 20 | US-09-534-845-23686 | Sequence 23686, A |
| 815 | 16 | 2.7 | 267 | 41 | US-60-092-036-925 | Sequence 925, App | 888 | 16 | 2.7 | 303 | 30 | US-09-770-175-3412 | Sequence 3412, Ap |
| 816 | 16 | 2.7 | 268 | 17 | US-09-304-517A-34074 | Sequence 34074, A | 889 | 16 | 2.7 | 305 | 40 | US-60-089-810-4044 | Sequence 4044, Ap |
| 817 | 16 | 2.7 | 268 | 17 | US-09-304-517A-35412 | Sequence 35412, A | 890 | 16 | 2.7 | 306 | 16 | US-09-275-082A-3588 | Sequence 3588, Ap |
| 818 | 16 | 2.7 | 268 | 17 | US-09-333-535A-15084 | Sequence 15084, A | 891 | 16 | 2.7 | 306 | 17 | US-09-304-517A-8891 | Sequence 8891, Ap |
| 819 | 16 | 2.7 | 268 | 17 | US-09-371-146A-34074 | Sequence 34074, A | 892 | 16 | 2.7 | 306 | 21 | US-09-371-146A-8891 | Sequence 8891, Ap |
| 820 | 16 | 2.7 | 268 | 17 | US-09-371-146A-55412 | Sequence 55412, A | 893 | 16 | 2.7 | 306 | 17 | US-09-540-229-100251 | Sequence 100251, A |
| 821 | 16 | 2.7 | 268 | 29 | US-09-724-866-13463 | Sequence 13463, A | 894 | 16 | 2.7 | 306 | 22 | US-09-585-800-486 | Sequence 486, App |
| 822 | 16 | 2.7 | 268 | 49 | US-09-724-866A-13463 | Sequence 13463, A | 895 | 16 | 2.7 | 306 | 32 | US-60-079-234-3588 | Sequence 3588, App |
| 823 | 16 | 2.7 | 268 | 49 | US-60-171-432-13463 | Sequence 13463, A | 896 | 16 | 2.7 | 307 | 19 | US-08-823-271-996 | Sequence 996, App |
| 824 | 16 | 2.7 | 269 | 21 | US-09-050-816-1388 | Sequence 1388, Ap | 897 | 16 | 2.7 | 307 | 17 | US-09-304-517A-40689 | Sequence 40689, A |
| 825 | 16 | 2.7 | 269 | 24 | US-09-540-229-56950 | Sequence 56950, A | 898 | 16 | 2.7 | 307 | 17 | US-09-371-146A-40689 | Sequence 40689, A |
| 826 | 16 | 2.7 | 269 | 24 | US-09-637-086A-18827 | Sequence 18827, A | 899 | 16 | 2.7 | 307 | 21 | US-09-540-764-25152 | Sequence 25152, A |
| 827 | 16 | 2.7 | 269 | 25 | US-09-654-617-243353 | Sequence 243353, A | 900 | 16 | 2.7 | 307 | 33 | US-60-015-513-996 | Sequence 996, App |
| 828 | 16 | 2.7 | 269 | 27 | US-09-684-016-243353 | Sequence 243353, A | 901 | 16 | 2.7 | 311 | 17 | US-09-321-209-3861 | Sequence 3861, Ap |
| 829 | 16 | 2.7 | 269 | 36 | US-60-043-982-1388 | Sequence 1388, Ap | 902 | 16 | 2.7 | 311 | 40 | US-60-086-856-3861 | Sequence 3861, Ap |
| 830 | 16 | 2.7 | 270 | 30 | US-09-770-175-1855 | Sequence 1855, Ap | 903 | 16 | 2.7 | 311 | 40 | US-60-087-762-188 | Sequence 188, App |
| 831 | 16 | 2.7 | 272 | 16 | US-09-244-000A-32805 | Sequence 32805, A | 904 | 16 | 2.7 | 312 | 41 | US-60-092-036-919 | Sequence 919, App |
| 832 | 16 | 2.7 | 272 | 17 | US-09-333-535A-9124 | Sequence 9124, Ap | 905 | 16 | 2.7 | 312 | 52 | US-60-207-458-40619 | Sequence 40619, A |
| 833 | 16 | 2.7 | 272 | 22 | US-09-552-086-23529 | Sequence 23529, A | 906 | 16 | 2.7 | 313 | 17 | US-09-304-517A-267775 | Sequence 267775, A |
| 834 | 16 | 2.7 | 272 | 27 | US-09-698-014-3457 | Sequence 3457, Ap | 907 | 16 | 2.7 | 313 | 17 | US-09-371-146A-26699 | Sequence 26699, A |
| 835 | 16 | 2.7 | 273 | 17 | US-09-504-517A-6861 | Sequence 6861, Ap | 908 | 16 | 2.7 | 313 | 22 | US-09-565-240-12252 | Sequence 12252, A |
| 836 | 16 | 2.7 | 273 | 21 | US-09-371-146A-6861 | Sequence 6861, Ap | 909 | 16 | 2.7 | 313 | 46 | US-60-141-134-4131 | Sequence 4131, Ap |
| 837 | 16 | 2.7 | 273 | 21 | US-09-540-212A-48720 | Sequence 48720, A | 910 | 16 | 2.7 | 315 | 16 | US-09-215-179A-6790 | Sequence 6790, Ap |
| 838 | 16 | 2.7 | 274 | 13 | US-08-917-047-1660 | Sequence 1660, Ap | 911 | 16 | 2.7 | 319 | 16 | US-09-270-849B-170818 | Sequence 170818, A |
| 839 | 16 | 2.7 | 274 | 21 | US-09-540-212A-9661 | Sequence 9661, Ap | 912 | 16 | 2.7 | 322 | 53 | US-60-217-080-18227 | Sequence 18227, A |
| 840 | 16 | 2.7 | 274 | 34 | US-60-025-203-1660 | Sequence 1660, Ap | 913 | 16 | 2.7 | 322 | 53 | US-09-411-999-31866 | Sequence 31866, A |
| 841 | 16 | 2.7 | 274 | 40 | US-60-087-762-186 | Sequence 186, App | 914 | 16 | 2.7 | 324 | 23 | US-09-605-700-20270 | Sequence 20270, Ap |
| 842 | 16 | 2.7 | 274 | 41 | US-60-092-036-923 | Sequence 923, App | 915 | 16 | 2.7 | 325 | 23 | US-09-605-700-3170 | Sequence 3170, Ap |
| 843 | 16 | 2.7 | 275 | 16 | US-09-628-858-1468 | Sequence 1468, Ap | 916 | 16 | 2.7 | 331 | 18 | US-09-411-999-33549 | Sequence 33549, A |
| 844 | 16 | 2.7 | 276 | 14 | US-09-244-000A-35826 | Sequence 35826, A | 917 | 16 | 2.7 | 331 | 29 | US-09-724-866-13442 | Sequence 13442, A |
| 845 | 16 | 2.7 | 276 | 18 | US-09-436-762A-25640 | Sequence 25640, A | 918 | 16 | 2.7 | 331 | 29 | US-09-724-866A-13442 | Sequence 13442, A |
| 846 | 16 | 2.7 | 279 | 16 | US-09-244-000A-19627 | Sequence 19627, A | 919 | 16 | 2.7 | 332 | 17 | US-60-171-432-13442 | Sequence 13442, A |
| 847 | 16 | 2.7 | 279 | 16 | US-09-244-000A-23588 | Sequence 23588, A | 920 | 16 | 2.7 | 332 | 17 | US-09-362-510-24841 | Sequence 24841, A |
| 848 | 16 | 2.7 | 279 | 16 | US-09-275-082A-1403 | Sequence 1403, Ap | 921 | 16 | 2.7 | 332 | 17 | US-09-362-510A-24841 | Sequence 24841, A |
| 849 | 16 | 2.7 | 279 | 17 | US-09-304-517A-290769 | Sequence 290769, A | 922 | 16 | 2.7 | 333 | 25 | US-09-654-617-305492 | Sequence 305492, A |

| Accession | Position | Sequence |
|-----------|----------|-------------------|
| C 923 | 16 | Sequence 305492 |
| C 924 | 16 | Sequence 6975, Ap |
| C 925 | 16 | Sequence 6975, Ap |
| C 926 | 16 | Sequence 6975, Ap |
| C 927 | 16 | Sequence 6975, Ap |
| C 928 | 16 | Sequence 6975, Ap |
| C 929 | 16 | Sequence 6975, Ap |
| C 930 | 16 | Sequence 6975, Ap |
| C 931 | 16 | Sequence 6975, Ap |
| C 932 | 16 | Sequence 6975, Ap |
| C 933 | 16 | Sequence 6975, Ap |
| C 934 | 16 | Sequence 6975, Ap |
| C 935 | 16 | Sequence 6975, Ap |
| C 936 | 16 | Sequence 6975, Ap |
| C 937 | 16 | Sequence 6975, Ap |
| C 938 | 16 | Sequence 6975, Ap |
| C 939 | 16 | Sequence 6975, Ap |
| C 940 | 16 | Sequence 6975, Ap |
| C 941 | 16 | Sequence 6975, Ap |
| C 942 | 16 | Sequence 6975, Ap |
| C 943 | 16 | Sequence 6975, Ap |
| C 944 | 16 | Sequence 6975, Ap |
| C 945 | 16 | Sequence 6975, Ap |
| C 946 | 16 | Sequence 6975, Ap |
| C 947 | 16 | Sequence 6975, Ap |
| C 948 | 16 | Sequence 6975, Ap |
| C 949 | 16 | Sequence 6975, Ap |
| C 950 | 16 | Sequence 6975, Ap |
| C 951 | 16 | Sequence 6975, Ap |
| C 952 | 16 | Sequence 6975, Ap |
| C 953 | 16 | Sequence 6975, Ap |
| C 954 | 16 | Sequence 6975, Ap |
| C 955 | 16 | Sequence 6975, Ap |
| C 956 | 16 | Sequence 6975, Ap |
| C 957 | 16 | Sequence 6975, Ap |
| C 958 | 16 | Sequence 6975, Ap |
| C 959 | 16 | Sequence 6975, Ap |
| C 960 | 16 | Sequence 6975, Ap |
| C 961 | 16 | Sequence 6975, Ap |
| C 962 | 16 | Sequence 6975, Ap |
| C 963 | 16 | Sequence 6975, Ap |
| C 964 | 16 | Sequence 6975, Ap |
| C 965 | 16 | Sequence 6975, Ap |
| C 966 | 16 | Sequence 6975, Ap |
| C 967 | 16 | Sequence 6975, Ap |
| C 968 | 16 | Sequence 6975, Ap |
| C 969 | 16 | Sequence 6975, Ap |
| C 970 | 16 | Sequence 6975, Ap |
| C 971 | 16 | Sequence 6975, Ap |
| C 972 | 16 | Sequence 6975, Ap |
| C 973 | 16 | Sequence 6975, Ap |
| C 974 | 16 | Sequence 6975, Ap |
| C 975 | 16 | Sequence 6975, Ap |
| C 976 | 16 | Sequence 6975, Ap |
| C 977 | 16 | Sequence 6975, Ap |
| C 978 | 16 | Sequence 6975, Ap |
| C 979 | 16 | Sequence 6975, Ap |
| C 980 | 16 | Sequence 6975, Ap |
| C 981 | 16 | Sequence 6975, Ap |
| C 982 | 16 | Sequence 6975, Ap |
| C 983 | 16 | Sequence 6975, Ap |
| C 984 | 16 | Sequence 6975, Ap |
| C 985 | 16 | Sequence 6975, Ap |
| C 986 | 16 | Sequence 6975, Ap |
| C 987 | 16 | Sequence 6975, Ap |
| C 988 | 16 | Sequence 6975, Ap |
| C 989 | 16 | Sequence 6975, Ap |
| C 990 | 16 | Sequence 6975, Ap |
| C 991 | 16 | Sequence 6975, Ap |
| C 992 | 16 | Sequence 6975, Ap |
| C 993 | 16 | Sequence 6975, Ap |
| C 994 | 16 | Sequence 6975, Ap |
| C 995 | 16 | Sequence 6975, Ap |

996 16 2.7 377 25 US-09-654-617-370906 Sequence 370906,
997 16 2.7 377 27 US-09-684-016-305492 Sequence 370906,
998 16 2.7 377 46 US-09-144-084-34443 Sequence 34443, A
999 16 2.7 377 46 US-60-145-146-39 Sequence 39, Appl
1000 16 2.7 378 19 US-09-528-409-10474 Sequence 10474, A

ALIGNMENTS

RESULT 1
US-08-988-242-1
Sequence 1, Application US/08988242
GENERAL INFORMATION:
APPLICANT: PARANOS-BACALA, GLAUCIA
APPLICANT: JOLIVER, MICHEL
APPLICANT: MANDRAND, BERNARD
TITLE OF INVENTION: NEW TRIPANOSOMA CRUZI ANTIGEN, GENE
TITLE OF INVENTION: ENCODING THEREFOR, AND METHODS OF DETECTING AND TREATING
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIVER & BERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/988,242
FILING DATE: 10-DEC-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: MPB 36400A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3402 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-988-242-1

Query Match 100.0%; Score 594; DB 13; Length 3402;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 594; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 caggtacagcgtacagcgttctgctcaatcgtacagcagcgtacgtgctgctgct 60
|||||
DB 1332 CAGGTACAGCCTACAGCCTTCTTCAATCTGACGACGACGCTACGCTGCTGCT 1231
|||||
QY 61 gacatgctgcatcgtatgacatctccatccgtcctccgcaaggagaagaacagcag 120
|||||
DB 1292 GACATGCTGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1351
|||||
QY 121 ccaggccaaacacacacgtgtagtgagcagcagcagcagcagcagcagcagcagc 180
|||||
DB 1352 CCAGGCCAAACACATGCTGATTCGACGACGACGACGACGACGACGACGACG 1411
|||||
QY 181 gacgagcagcgtacacgtacacgtacacgtacacgtacacgtacacgtacacgt 240
|||||

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

| | | | | |
|-----------------------|-----------------|---------------|-----------|--------------|
| Query Match | 100.0%; | Score 594; | DB 15; | Length 3402; |
| Best Local Similarity | 100.0%; | Pred. NO. 0; | | |
| Matches 594; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |

[illegible]

```

1      RESULT 3
2      US-08-988-242-5
3      ; Sequence 5, Application US/08988242
4      ; GENERAL INFORMATION:
5      ; APPLICANT: PARANHOS-BACCALA, GLAUCIA
6      ; APPLICANT: LESENECHAL, MYLENE
7      ; APPLICANT: JOLIVET, MICHEL
8      ; APPLICANT: MANDRAND, BERNARD
9      ; TITLE OF INVENTION: NEW TYPANOSOMA CRUZI ANTIGEN, GENE
10     ; TITLE OF INVENTION: ENCODING THEREFOR, AND METHODS OF DETECTING AND TREATINGS
11     ; TITLE OF INVENTION: CHAGAS DISEASE
12     ; NUMBER OF SEQUENCES: '19
13     ; CORRESPONDENCE ADDRESS:
14     ; ADDRESSEE: OLIVÉ & BERRIDGE, PLC
15     ; STREET: P.O. BOX 19928
16     ; CITY: Alexandria
17     ; STATE: Virginia
18     ; COUNTRY: U.S.A.
19     ; ZIP: 22320
20     ; COMPUTER READABLE FORM:
21     ; MEDIUM TYPE: Floppy disk
22     ; COMPUTER: IBM PC compatible
23     ; OPERATING SYSTEM: PC-DOS/MS-DOS
24     ; SOFTWARE: PatentIn Release #1.0, Version #1.30
25     ; CURRENT APPLICATION DATA:
26     ; APPLICATION NUMBER: US/08/988,242
27     ; FILING DATE: 10-DEC-1997
28     ;

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: MPB 36400A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-988-242-5

Query Match 3.7%; Score 22; DB 13; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.93;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 cagccgacgctgctgcctcct 56
|||||
Db 1 CAGCCGACGCTAGCTGCTCT 22

RESULT 4
US-09-138-736-9
Sequence 9, Application US/09138736
GENERAL INFORMATION:
APPLICANT: PARANHOS-BACCALA, Glaucia
APPLICANT: LESENECHAL, Mylene
TITLE OF INVENTION: NEW TRYPARANOSOMA CRUZI ANTIGEN, AND GENE
TITLE OF INVENTION: ENCODING THE LATTER, THEIR APPLICATION TO THE DETECTION OF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Oliff & Berridge
STREET: 700 South Washington Street, Suite 300
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/138,736
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/480,917
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: MPB 36400
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-138-736-9

Query Match 3.7%; Score 22; DB 15; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.93;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 cagccgacgctgctgcctcct 56
|||||
Db 1 CAGCCGACGCTAGCTGCTCT 22

RESULT 5
US-09-654-617-126156
Sequence 126156, Application US/09654617
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Liu, Jindong
TITLE OF INVENTION: Annotated Plant Genes
FILE REFERENCE: 38-21(15097)D
CURRENT APPLICATION NUMBER: US/09/654,617
CURRENT FILING DATE: 2000-09-05
NUMBER OF SEQ ID NOS: 463173
SEQ ID NO 126156
LENGTH: 718
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-654-617-126156

Query Match 3.4%; Score 20; DB 25; Length 718;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 ctctgcgcgtgctgcatcc 231
|||||
Db 586 ctctgcgcgtgctgcatcc 605

RESULT 6
US-09-684-016-126156
Sequence 126156, Application US/09684016
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Liu, Jindong
TITLE OF INVENTION: Annotated Plant Genes
FILE REFERENCE: 38-21(15097)D
CURRENT APPLICATION NUMBER: US/09/684,016
CURRENT FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: US 09/654,617
PRIOR FILING DATE: 2000-09-05
NUMBER OF SEQ ID NOS: 463173
SEQ ID NO 126156
LENGTH: 718
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-684-016-126156

Query Match 3.4%; Score 20; DB 27; Length 718;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 ctctgcgcgtgctgcatcc 231
|||||
Db 586 ctctgcgcgtgctgcatcc 605

RESULT 7
US-09-620-392-64793/C
Sequence 64793, Application US/09620392
GENERAL INFORMATION:
APPLICANT: Boukharov, Andrey A.
APPLICANT: Kovalic, David K.
APPLICANT: Liu, Jindong

```

; APPLICANT: McIninch, James
; TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof
; FILE REFERENCE: 38-21(51237)E
; CURRENT APPLICATION NUMBER: US/09/620,392
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 69652
; SEQ ID NO 64793
; LENGTH: 5679
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-620-392-64793
```

```

Query Match
Best Local Similarity 100.0%; Pred. No. 12; Length 5679;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 215 ctgcgcgtcgtcatccct 234
|||||
Db 3043 CTGCCGCTGCTGCATCCCT 3024
```

```

RESULT 8
US-09-702-134-21010/c
; Sequence 21010, Application US/09702134
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Plant Genome Sequence and Uses Thereof
; FILE REFERENCE: 38-21(51237)F
; CURRENT APPLICATION NUMBER: US/09/702,134
; CURRENT FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 52202
; SEQ ID NO 21010
; LENGTH: 5679
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-702-134-21010
```

```

Query Match
Best Local Similarity 100.0%; Pred. No. 12; Length 5679;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 215 ctgcgcgtcgtcatccct 234
|||||
Db 3043 CTGCCGCTGCTGCATCCCT 3024
```

```

RESULT 9
US-60-173-464-11355
; Sequence 11355, Application US/60173464
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W.D.
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
; FILE REFERENCE: C1000173
; CURRENT APPLICATION NUMBER: US/60/173,464
; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 30269
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11355
; LENGTH: 11335
; TYPE: DNA
; ORGANISM: Drosophila
US-60-173-464-11355
```

```

Query Match
Best Local Similarity 100.0%; Pred. No. 12; Length 11335;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 50 ggcgtcgtcgtgacatgctcg 69
|||||
Db 6835 ggcgtcgtcgtgacatgctcg 6854
```

```

RESULT 10
US-60-191-637-13953/c
; Sequence 13953, Application US/60191637
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; FILE REFERENCE: C1000392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13953
; LENGTH: 11335
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-60-191-637-13953
```

```

Query Match
Best Local Similarity 100.0%; Pred. No. 12; Length 11335;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 50 ggcgtcgtcgtgacatgctcg 69
|||||
Db 4501 GCGTCTGCTGACATGTCG 4482
```

```

RESULT 11
US-60-191-681-10986/c
; Sequence 10986, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; FILE REFERENCE: C1000390
; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 30973
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 10986
; LENGTH: 11335
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-60-191-681-10986
```

```

Query Match
Best Local Similarity 100.0%; Pred. No. 12; Length 11335;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 50 ggcgtcgtcgtgacatgctcg 69
|||||
Db 4501 GCGTCTGCTGACATGTCG 4482
```

```

RESULT 12
US-60-167-217-13998
; Sequence 13998, Application US/60167217
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W. D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
```

```

:
: TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES
:
: TITLE OF INVENTION: THERMOF
:
: FILE REFERENCE: CLO000152
:
: CURRENT APPLICATION NUMBER: US/60/167,217
:
: CURRENT FILING DATE: 1999-11-24
:
: NUMBER OF SEQ ID NOS: 23195
:
: SOFTWARE: FastSeq for Windows Version 4.0
:
: SEQ ID NO 13398
:
: LENGTH: 11338
:
: TYPE: DNA
:
: ORGANISM: Drosophila
:
: US-60-167-217-13998

```

| | | | | |
|-----------------------|----------------|-----------------------|----------|--------------|
| Query Match | 3.4% | Score 20 | DB 48 | Length 11338 |
| Best Local Similarity | 100.0% | Pred. No. 12 | | |
| Matches 20 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 |
| QY | 50 | gcgcctcgcgcgcacatgcgc | 69 | |
| | | | | |
| db | 6836 | gcgcctcgcgcgcacatgcgc | 6857 | |

```

RESULT 13
US-09-534-859-19/c
: Sequence 19, Application US/09534859
GENERAL INFORMATION:
APPLICANT: Bush, David F.
APPLICANT: Last, Robert L.
APPLICANT: Levin, Irena M.
APPLICANT: Norris, Susan R.
APPLICANT: Parnell, Laurence D.
APPLICANT: Rounsley, Steven D.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: PLANT POLYMORPHIC MARKERS AND USES THEREOF
FILE REFERENCE: 38-10(115493)B
CURRENT APPLICATION NUMBER: US/09/534,859
CURRENT FILING DATE: 2000-03-29
NUMBER OF SEQ ID NOS: 1127
SEQ ID NO 19
LENGTH: 91470
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-534-859-19

```

```

Query Match          3.4%; Score 20; DB 20; Length 91470;
Best Local Similarity 100.0%; Pred. NO. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
OY 212 ctctgcgcgtcgtcgtccatcc 231
    |||||
db 42747 ctttgcgcgcgtcgtccatcc 42728

```

```

US-RESULT 14
US-09-528-237A-1285
: Sequence 1285, Application US/09528237A
: GENERAL INFORMATION:
: APPLICANT: Venter, J. Craig
: TITLE OF INVENTION: The Drosophila Genome, Primary Nucleic
: TITLE OF INVENTION: Acid Sequences, Systems Containing The Nucleic Acid
: TITLE OF INVENTION: Sequences and Uses Thereof
: FILE REFERENCE: CL000284
: CURRENT APPLICATION NUMBER: US/09/528,237A
: CURRENT FILING DATE: 2000-03-17
: NUMBER OF SEQ. ID NOS: 2926
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ. ID NO 1285
: LENGTH: 226706
: TYPE: DNA
: ORGANISM: Drosophila
: US-09-528-237A-1285

```

| | | | | |
|-----------------------|--------|----------------------|--------|----------------|
| Query Match | 3.4% | Score 20; | DB 19; | Length 226706; |
| Best Local Similarity | 100.0% | Pred. No. 12; | | |
| Matches | 20; | Conservative | 0; | Mismatches 0; |
| | | | | Indels 0; |
| | | | | Gaps 0; |
| QY | 50 | gcgcctgcgtacatgctcg | 69 | |
| | | | | |
| Db | 213597 | gcgcctcgtgcacatgctcg | 213616 | |

QY 50 gcgtcctggtcatgtcg 69
|||||
Db 213597 gcgtcctggtcatgtcg 213616

```

RESULT 15
US-60-182-316-2055/c
: Sequence 2055, Application US/60182316
: GENERAL INFORMATION:
: APPLICANT: Curtlis, Anne L.
: APPLICANT: Lagace, Robert E.
: APPLICANT: Klingler, Tod M.
: APPLICANT: Stuve, Laura L.
: TITLE OF INVENTION: CPG Island Polynucleotides
: FILE REFERENCE: PX-0003 P
: CURRENT APPLICATION NUMBER: US/60/182,316
: CURRENT FILING DATE: 2000-02-10
: NUMBER OF SEQ ID NOS: 14,630
: SOFTWARE: PERL Program
: SEQ ID NO 2055
:
: LENGTH: 718
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: CPG-991027_B15_masked_fa.Contig49773
US-60-182-316-2055

```

| | | | | | | |
|----|-----|-----------------------|--------|--------------|-------|---------------------------------|
| | | Query Match | 3.2% | Score 19 | DB 50 | Length 718 |
| | | Best Local Similarity | 100.0% | Pred. No. | 41 | |
| | | Matches | 19 | Conservative | 0 | Mismatches 0; Indels 0; Gaps 0; |
| OY | 280 | gccgcgacgcagatcgactg | 298 | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| db | 213 | gccgcggcgcgcgatccgtc | 195 | | | |

```

RESULT 16
US-09-620-392-32691
: Sequence 32691, Application US/09620392
: GENERAL INFORMATION:-
: APPLICANT: Boukharov, Andrey A.
: APPLICANT: Kovalic, David K.
: APPLICANT: Liu, Jingdong
: APPLICANT: McIninch, James
: TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof
: FILE REFERENCE: 38-21(51237)E
: CURRENT APPLICATION NUMBER: US/09/620,392
: CURRENT FILING DATE: 2000-07-19
: NUMBER OF SEQ ID NOS: 69652
: SEQ ID NO 32691
: LENGTH: 9855
: TYPE: DNA
: ORGANISM: Oryza sativa
US-09-620-392-32691

```

| | | | | |
|-----------------------|----------------|--------------------|----------|-------------|
| Query Match | 3.2% | Score 19 | DB 24 | Length 9855 |
| Best Local Similarity | 100.0% | Pred. No. 42 | | |
| Matches 19 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 |
| OY | 387 | gccgcgcacgacgagctc | 405 | |
| | | | | |
| | | | | |
| | | | | |
| Db | 259 | gccgcgcacgacgagctc | 277 | |

```
RESULT 17
US-09-620-392-67713
; Sequence 67713, Application US/09620392
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Kovacic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: McIninch, James
; TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof
; FILE REFERENCE: 38-21(51237)E
; CURRENT APPLICATION NUMBER: US/09/620,392
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 69652
; SEQ ID NO 67713
; LENGTH: 9895
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-620-392-67713

Query Match          3.2%; Score 19; DB 24; Length 9895;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 gccggccacgacgagctct 405
    |||
Db 8100 gccggccacgacgagctct 8118

RESULT 18
US-09-407-562-32/c
; Sequence 32, Application US/09407562
; GENERAL INFORMATION:
; APPLICANT: Kathryn Meek
; TITLE OF INVENTION: Genetic Test For Equine Severe
; TITLE OF INVENTION: Combined Immunodeficiency Disease
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Benjamin A. Adler
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/407,562
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/970,269
; FILING DATE: November 14, 1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler Ph.D., Benjamin A.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5860
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-777-2321
; TELEFAX: 713-777-6908
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11873 bp
; TYPE: nucleic acid
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: other nucleic acid
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; FEATURE:
US-09-407-562-32
```

```
ANTI-SENSE: no
ORIGINAL SOURCE:
FEATURE:
US-09-407-562-32

Query Match          3.2%; Score 19; DB 18; Length 11873;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 actctgcgcgtctgcac 229
    |||
Db 6446 actctgcgcgtctgcac 6428

RESULT 19
US-09-407-562-31/c
; Sequence 31, Application US/09407562
; GENERAL INFORMATION:
; APPLICANT: Kathryn Meek
; TITLE OF INVENTION: Genetic Test For Equine Severe
; TITLE OF INVENTION: Combined Immunodeficiency Disease
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Benjamin A. Adler
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/407,562
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/970,269
; FILING DATE: November 14, 1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler Ph.D., Benjamin A.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5860
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-777-2321
; TELEFAX: 713-777-6908
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11878 bp
; TYPE: nucleic acid
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: other nucleic acid
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; FEATURE:
US-09-407-562-31

Query Match          3.2%; Score 19; DB 18; Length 11878;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 actctgcgcgtctgcac 229
    |||
Db 6446 actctgcgcgtctgcac 6428
```


| | Matches | 19; Conservative | 0; Mismatches | 0; Indels | 0; Gaps |
|----|---------|--------------------|---------------|-----------|---------|
| QY | 387 | gccggcagacagagttc | 405 | | |
| | | | | | |
| Db | 9468 | gccggcagacagaggtct | 9450 | | |

```

RESULT 24
US-09-702-134-5301
: Sequence 5301, Application US/09702134
GENERAL INFORMATION:
APPLICANT: Boukharov, Andrey A.
APPLICANT: Cao, Yongwei
APPLICANT: Kovacic, David K.
APPLICANT: Liu, Jingdong
APPLICANT: McIninch, James
APPLICANT: Wu, Wei
: TITLE OF INVENTION: Plant Genome Sequence and Uses Thereof
: FILE REFERENCE: 38-21(51237)F
: CURRENT APPLICATION NUMBER: US/09/702,134
: CURRENT FILING DATE: 2000-10-31
: NUMBER OF SEQ ID NOS: 52202
: SEQ ID NO 5301
: LENGTH: 155989
: TYPE: DNA
: ORGANISM: Oryza sativa
US-09-702-134-5301

```

| | | | | |
|--------------------------|--------|---------------|--------|----------------|
| Query Match | 3.2% | Score 19: | DB 28; | Length 159989; |
| Best Local Similarity | 100.0% | Pred. No. 42; | | |
| Matches 19; Conservative | 0; | Mismatches | 0; | Gaps 0 |

RESULT 25
 US-08-988-242-4/c
 : Sequence 4, Application US/08988242
 : GENERAL INFORMATION:
 : APPLICANT: PARANHOS-BACCALA, GLAUCIA
 : APPLICANT: LESENECHAL, MYLENE
 : APPLICANT: JOLIVET, MICHEL
 : APPLICANT: MANDRAND, BERNARD
 : TITLE OF INVENTION: NEW TRYPANOSOMA CRUZI ANTIGEN, GENE
 : TITLE OF INVENTION: ENCODING THEREFOR, AND METHODS OF DETECTING AND TREATING
 : TITLE OF INVENTION: CHAGAS DISEASE
 : NUMBER OF SEQUENCES: 19
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: OLIFF & BERRIDGE, PLC
 : STREET: P.O. BOX 19928
 : CITY: Alexandria
 : STATE: Virginia
 : COUNTRY: U.S.A.
 : ZIP: 22320
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/988,242
 : FILING DATE: 10-DEC-1997
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Berridge, William P.
 : REGISTRATION NUMBER: 30,024
 : REFERENCE/DOCKET NUMBER: WPB 36400A
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 703-836-6400
 : TELEFAX: 703-836-2787

```

; INFORMATION FOR SEQ ID NO: 4:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 18 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: DNA (genomic)
;
US-08-988-242-4

```

| | | | | |
|-----------------------|-----------------|----------------------|-----------|------------|
| Query Match | 3.0%; | Score 18; | DB 13; | Length 18; |
| Best Local Similarity | 100.0%; | Pred. No. 1.4e+02; | | |
| Matches 18; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Oy | 211 | actctgcgcgtcgtcgca | 228 | |
| | | | | |
| Db | 18 | actctgcgcgcgtcgtcgca | 1 | |

```

RESULT 26
US-09-138-736-5
; Sequence 5, Application US/09138736
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: LESENECHAL, Mylene
; APPLICANT: JOLIVET, Michel
; TITLE OF INVENTION: NEW TRYPAOSOMA CRUZI ANTIGEN, AND GENE
; TITLE OF INVENTION: ENCODING THE LATTER, THEIR APPLICATION TO THE DETECTION OF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Bertridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/138,736
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,917
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bertridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-138-736-5

Query Match 3.0%; Score 18; DB 15; Length 18;
Best Local Similarity 100.0%; Pred.No.1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 tcgggcactgaacgcggcg 189
|||||
Db 1 tcgggcactgaacgcggcg 18

```

```
RESULT 27
US-09-138-736-8/c
; Sequence 8, Application US/09138736
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: LESENECHAL, Mylene
; APPLICANT: JOLIVET, Michel
; TITLE OF INVENTION: NEW TRYPAOSOMA CRUZI ANTIGEN, AND GENE
; TITLE OF INVENTION: ENCODING THE LATTER; THEIR APPLICATION TO THE DETECTION OF
; TITLE OF INVENTION: CHAGAS DISEASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/138,736
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,917
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: MPB 36400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-138-736-8

Query Match          3.08; Score 18; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 actctgcgcgtctgca 228
      |||||||
Db 18 ACTTCTGCCGCTGCTGCA 1

RESULT 28
US-09-366-691A-1504/c
; Sequence 1504, Application US/09366691A
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A., et al.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 38
; FILE REFERENCE: PO-38
; CURRENT APPLICATION NUMBER: US/09/366,691A
; CURRENT FILING DATE: 1999-08-04
; PRIOR APPLICATION NUMBER: 60/095,485
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 8319
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1504
; LENGTH: 70
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (60)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (65)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (70)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-366-691A-1504

Query Match          3.08; Score 18; DB 17; Length 70;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 tgcgcgtcgtcgtcctccc 233
      |||||||
Db 31 TGCCGCTGCTGCATCCCTCC 14

RESULT 29
US-09-076-667-928/c
; Sequence 928, Application US/09076667
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerbloom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; NUMBER OF SEQUENCES: 4483
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/076,667
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 928:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 85 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; IMMEDIATE SOURCE:
```


CLONE: 3244053H1
US-09-076-667-928

Query Match
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 agaacagcagccagccca 128
Db 29 AGAACAGCAGCCAGCCCA 12

RESULT 30
US-09-540-229-15688/C
; Sequence 15688, Application US/09540229
; GENERAL INFORMATION:
; APPLICANT: Sellhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF NERVOUS SYSTEM AND SENSORY ORGANS
; FILE REFERENCE: PD-1033 CIP
; CURRENT APPLICATION NUMBER: US/09/540,229
; CURRENT FILING DATE: 2000-03-31
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 193582
; SOFTWARE: PERL Program
; SEQ ID NO 15688
; LENGTH: 85
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00290743
; NAME/KEY: unsure
; LOCATION: 3, 81
; OTHER INFORMATION: a, t, c, g, or other
US-09-540-229-15688

Query Match
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 agaacagcagccagccca 128
Db 29 AGAACAGCAGCCAGCCCA 12

RESULT 31
US-60-048-002-928/C
; Sequence 928, Application US/60048002
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN BRAIN
; NUMBER OF SEQUENCES: 4483
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA

ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: word perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/048,002
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PD-0370P
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 928:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
IMMEDIATE SOURCE:
CLONE: 3244053H1
US-60-048-002-928

Query Match
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 agaacagcagccagccca 128
Db 29 AGAACAGCAGCCAGCCCA 12

RESULT 32
US-09-513-991-2039
; Sequence 2039, Application US/09513991
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Fraser, Christopher C.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Jakubowski, Joseph A.
; TITLE OF INVENTION: NOCLEIC ACID MOLECULES DERIVED FROM
; FILE REFERENCE: 1600.1083-001
; CURRENT APPLICATION NUMBER: US/09/513,991
; EARLIER APPLICATION NUMBER: 60/123,397
; EARLIER FILING DATE: 1999-03-08
; NUMBER OF SEQ ID NOS: 3378
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2039
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)-(180)
; OTHER INFORMATION: n = A,T,C or G
US-09-513-991-2039

Query Match
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 tccggtctctccgcagcg 109
Db 1 tccggtctctccgcagcg 18

RESULT 33
US-08-810-326-2151/C
; Sequence 2151, Application US/08810326
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angela E.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN LUNG
; NUMBER OF SEQUENCES: 3314
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,326
; FILING DATE: HEREWITH
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/012,699
; FILING DATE: FEBRUARY 29, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/015,173
; FILING DATE: APRIL 10, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CERBONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PD-0121 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 2151:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 765907
; US-08-810-326-2151

Query Match 3.0%; Score 18; DB 12; Length 234;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 gtagctgcgtcgtcgctg 61
|||||
Db 136 gtactgcgtcgtcgtcg 119

RESULT 34
US-09-540-212A-11685/C
; Sequence 11685, Application US/09540212A
; GENERAL INFORMATION:

; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF AIRWAY AND LUNG SYSTEM TISSUE
; FILE REFERENCE: PD-1034 CIP
; CURRENT APPLICATION NUMBER: US/09/540,212A
; CURRENT FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 67551
; SOFTWARE: PERL Program
; SEQ ID NO 11685
; LENGTH: 234
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00594097
; LOCATION: 84, 99
; OTHER INFORMATION: a, t, c, g, or other
; US-09-540-212A-11685

Query Match 3.0%; Score 18; DB 21; Length 234;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 gtagctgcgtcgtcgctg 61
|||||
Db 136 gtactgcgtcgtcgtcg 119

RESULT 35
US-08-878-507-222/C
; Sequence 222, Application US/08878507
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angela E.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN LUNG
; NUMBER OF SEQUENCES: 1429
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,507
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/016,092
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: GLAISTER, DEBRA J., PATENT AGENT
; REGISTRATION NUMBER: 33888
; REFERENCE/DOCKET NUMBER: PD-0188P
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 222:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 1360155
US-08-878-507-222

Query Match 3.0%: Score 18; DB 12; Length 237;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 gtagctgcgtcctgctg 61
|||||
Db 165 gtagctgcgtcctgctg 148

RESULT 36
US-09-540-212A-16821/C
Sequence 16821, Application US/09540212A
GENERAL INFORMATION:
APPLICANT: Seilhamer, Jeffrey J.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullahy, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES OF AIRWAY AND LUNG SYSTEM TISSUE
FILE REFERENCE: PD-1034 CIP
CURRENT APPLICATION NUMBER: US/09/540.212A
CURRENT FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 67551
SOFTWARE: PERL Program
SEQ ID NO 16821
LENGTH: 237
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: hu00152336
US-09-540-212A-16821

Query Match 3.0%: Score 18; DB 21; Length 237;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 gtagctgcgtcctgctg 61
|||||
Db 165 gtagctgcgtcctgctg 148

RESULT 37
US-60-016-092-222/C
Sequence 222, Application US/60016092
GENERAL INFORMATION:
APPLICANT: Gooding, Douglas H.
APPLICANT: Stuve, Laura L.
APPLICANT: Stuart, Susan G.
APPLICANT: Ito, Laura Y.
APPLICANT: Akherblom, Ingrid E.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Naughton, Rebecca E.
APPLICANT: Klingler, Tod M.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
HUMAN LUNG
NUMBER OF SEQUENCES: 1429

CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/016.092
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: GLAISTER, DEBRA J., PATENT AGENT
REGISTRATION NUMBER: 33888
REFERENCE/DOCKET NUMBER: PD-0188P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 222:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 1360155
US-60-016-092-222

Query Match 3.0%: Score 18; DB 33; Length 237;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 gtagctgcgtcctgctg 61
|||||
Db 165 gtagctgcgtcctgctg 148

RESULT 38
US-09-100-454-2483/C
Sequence 2483, Application US/09100454
GENERAL INFORMATION:
APPLICANT: Gooding, Douglas H.
APPLICANT: Stuart, Susan G.
APPLICANT: Ito, Laura Y.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
HUMAN LUNG
NUMBER OF SEQUENCES: 3617
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100.454
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/051.749

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PD-0388P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 2483:
SEQUENCE CHARACTERISTICS:
LENGTH: 248 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 2691009H1
US-09-100-454-2483

Query Match 3.0%; Score 18; DB 15; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 gtagctgcgtcctgcgtcg 61
|||||

DB 209 GTAGCTGCGTCTGGCTG 192

RESULT 39

US-09-540-212A-35497/C
Sequence 35497, Application US/09540212A
GENERAL INFORMATION:
APPLICANT: Sellhauer, Jeffrey J.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullahy, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES OF AIRWAY AND LUNG SYSTEM TISSUE
FILE REFERENCE: PD-1034 CIP
CURRENT APPLICATION NUMBER: US/09/540,212A
CURRENT FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 67551
SOFTWARE: PERL Program
SEQ ID NO 35497
LENGTH: 248
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: hu00026037
US-09-540-212A-35497

Query Match 3.0%; Score 18; DB 21; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 gtagctgcgtcctgcgtcg 61
|||||

DB 209 GTAGCTGCGTCTGGCTG 192

RESULT 40

US-60-051-749-2483/C
Sequence 2483, Application US/60051749
GENERAL INFORMATION:
APPLICANT: Gooding, Douglas H.
APPLICANT: Stuart, Susan G.
APPLICANT: Ito, Laura Y.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM

TITLE OF INVENTION: HUMAN LUNG
NUMBER OF SEQUENCES: 3617
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/051,749
FILING DATE: HERewith

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PD-0388P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166

INFORMATION FOR SEQ ID NO: 2483:

SEQUENCE CHARACTERISTICS:
LENGTH: 248 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 2691009H1
US-60-051-749-2483

Query Match 3.0%; Score 18; DB 37; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 gtagctgcgtcctgcgtcg 61
|||||

DB 209 GTAGCTGCGTCTGGCTG 192

RESULT 41

US-60-048-722-3053/C
Sequence 3053, Application US/60048722
GENERAL INFORMATION:
APPLICANT: Gooding, Douglas H.
APPLICANT: Stuve, Laura L.
APPLICANT: Stuart, Susan G.
APPLICANT: Ito, Laura Y.
APPLICANT: Akerblom, Ingrid E.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Naughton, Rebecca E.
APPLICANT: Klingler, Tod M.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
NUMBER OF SEQUENCES: 3106
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/60/048,722
;; FILING DATE: HEREMITH
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: CERRONE, MICHAEL C.
;; REGISTRATION NUMBER: 39,132
;; REFERENCE/DOCKET NUMBER: PD-0381P
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 855-0555
;; TELEFAX: (415) 845-4166
;; INFORMATION FOR SEQ ID NO: 3053:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 249 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; IMMEDIATE SOURCE:
;; CLONE: 2933007H1
;; US-60-048-722-3053

Query Match 3.0%; Score 18; DB 36; Length 249;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 104 gcagggaagacagcagc 121
|||||
Db 190 GCAGGAGAGACAGCAGC 173

RESULT 42

US-08-992-625-3252/C

; Sequence 3252, Application US/08992625

; GENERAL INFORMATION:

; APPLICANT: Gooding, Douglas H.

; APPLICANT: Stuve, Laura L.

; APPLICANT: Stuart, Susan G.

; APPLICANT: Ito, Laura Y.

; APPLICANT: Akerblom, Ingrid E.

; APPLICANT: Delegeane, Angelo M.

; APPLICANT: Naughton, Rebecca E.

; APPLICANT: Klingler, Tod M.

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM

; NUMBER OF SEQUENCES: 3446

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/992,625

; FILING DATE: HEREMITH

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/036,187

; FILING DATE: DECEMBER 18, 1996

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: CERRONE, MICHAEL C.

; REGISTRATION NUMBER: 39,132

; REFERENCE/DOCKET NUMBER: PD-0288 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 855-0555

;; TELEFAX: (650) 845-4166
;; INFORMATION FOR SEQ ID NO: 3252:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 253 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; IMMEDIATE SOURCE:
;; CLONE: 2264459H1
;; US-08-992-625-3252

Query Match 3.0%; Score 18; DB 13; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 gtagctgcgtcctgcgtg 61
|||||
Db 59 GTAGCTGCCTCCTGCCTG 42

RESULT 43

US-09-540-764-10498/C

; Sequence 10498, Application US/09540764

; GENERAL INFORMATION:

; APPLICANT: Sellhammer, Jeffrey J.

; APPLICANT: Delegeane, Angelo M.

; APPLICANT: Stuart, Susan G.

; APPLICANT: Stuve, Laura L.

; APPLICANT: Mullahy, Sara J.

; APPLICANT: Naughton, Rebecca E.

; TITLE OF INVENTION: POLYNUCLEOTIDES OF CELLS AND TISSUES OF THE FEMALE REPRODUCTI

; FILE REFERENCE: PD-1028 CIP

; CURRENT APPLICATION NUMBER: US/09/540,764

; CURRENT FILING DATE: 2000-03-30

; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 61458

; SOFTWARE: PERL Program

; SEQ ID NO 10498

; LENGTH: 253

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; OTHER INFORMATION: Incyte ID No: hu08822021

; US-09-540-764-10498

Query Match 3.0%; Score 18; DB 21; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 gtagctgcgtcctgcgtg 61
|||||
Db 59 GTAGCTGCCTCCTGCCTG 42

RESULT 44

US-60-197-872-21718

; Sequence 21718, Application US/60197872

; GENERAL INFORMATION:

; APPLICANT: Bougri, Olegs

; APPLICANT: Byrum, Joseph R.

; APPLICANT: De La Pena, Robert C.

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Shukla, Hridayabhiranjan

; TITLE OF INVENTION: Nucleic acid Molecules and other molecules associated w1

; FILE REFERENCE: 38-21(51892)A

; CURRENT APPLICATION NUMBER: US/60/197,872

; CURRENT FILING DATE: 2000-04-19

; NUMBER OF SEQ ID NOS: 76255

SEQ ID NO 21718
LENGTH: 260
TYPE: DNA
ORGANISM: Oryza sativa nipponbare
OTHER INFORMATION: Clone ID: LIB3433-013-Q6-K1-F11
US-60-197-872-21718

Query Match 3.0%; Score 18; DB 51; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 394 acgacgagctctacgcgcg 411
|||||
DB 107 acgacgagctctacgcgcg 124

RESULT 45
US-08-104-507A-4311/C
Sequence 4311, Application US/08104507A

GENERAL INFORMATION:
APPLICANT: Haseltine, William A.
APPLICANT: Rosen, Craig A.
APPLICANT: Ruben, Steve
APPLICANT: Dillon, Patrick
APPLICANT: Li, Haodong
APPLICANT: Adams, Mark D.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products
NUMBER OF SEQUENCES: 7483
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart, & Olstein
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: New Jersey
COUNTRY: USA
ZIP: 07068

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 5.0
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/104, 507A

CLASSIFICATION:
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Herron, Charles J.
REGISTRATION NUMBER: 28,019
REFERENCE/DOCKET NUMBER: 325800-37
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 994-1700
TELEFAX: (201) 994-1744
INFORMATION FOR SEQ ID NO: 4311:

SEQUENCE CHARACTERISTICS:
LENGTH: 261 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-104-507A-4311

Query Match 3.0%; Score 18; DB 5; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 gtacgtcgctctgcgcg 61
|||||
DB 18 ctacgtcgctctgcgcg 1

RESULT 46
US-08-104-507C-4311/C
Sequence 4311, Application US/08104507C

GENERAL INFORMATION:
APPLICANT: Adams, Mark D.
APPLICANT: Kelley, Jenny
APPLICANT: Haseltine, William A.
APPLICANT: Rosen, Craig A.
APPLICANT: Ruben, Steve
APPLICANT: Dillon, Patrick J.
TITLE OF INVENTION: Human Genes, Sequences and Expression
NUMBER OF SEQUENCES: 7485
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/104, 507C
FILING DATE: August 9, 1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PO-04
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 4311:
SEQUENCE CHARACTERISTICS:
LENGTH: 261 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-104-507C-4311

Query Match 3.0%; Score 18; DB 5; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 gtacgtcgctctgcgcg 61
|||||
DB 18 gtacgtcgctctgcgcg 1

RESULT 47
US-09-270-849B-53924/C
Sequence 53924, Application US/09270849B

GENERAL INFORMATION:
APPLICANT: Swimmer et al.
TITLE OF INVENTION: Insect genome survey devices
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/270, 849B
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 195450
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 53924
LENGTH: 269
TYPE: DNA
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-270-849B-553924

Query Match 3.0%; Score 18; DB 16; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 240 accccctgttcagcgcc 257
|||||
DB 25 ACCCCCTGTTCAGCGCC 8

RESULT 48
US-09-298-329A-4053/C
Sequence 4053, Application US/09298329A
GENERAL INFORMATION:
APPLICANT: Laljudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
FILE REFERENCE: PL-0012 US
CURRENT APPLICATION NUMBER: US/09/298,329A
CURRENT FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/085,331
PRIOR FILING DATE: 1998-05-12
NUMBER OF SEQ ID NOS: 6332
SOFTWARE: PERL Program
SEQ ID NO 4053
LENGTH: 271
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: 700454335H1
NAME/KEY: unsure
LOCATION: 56, 83, 164, 235
OTHER INFORMATION: a, t, c, g, or other
US-09-298-329A-4053

Query Match 3.0%; Score 18; DB 16; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 524 ccacagccgagcggaac 541
|||||
DB 102 CCACAGCCGCGGCGAAC 85

RESULT 49
US-60-087-762-211/C
Sequence 211, Application US/60087762
GENERAL INFORMATION:
APPLICANT: Christine Shemmaker
APPLICANT: Dang Yang Ke
APPLICANT: Jingdong Liu
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules
TITLE OF INVENTION: Associated with the Isoflavone pathway
NUMBER OF SEQUENCES: 336
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lawrence M. Lavin, Jr.
STREET: BB4F
STREET: Monsanto Company
STREET: 700 Chesterfield Parkway North
CITY: St. Louis
STATE: MO
COUNTRY: United States
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette (3.5in, 1.44 Mb)
COMPUTER: IBM PC/XT/AT, IBM PS/2 or compatibles

OPERATING SYSTEM: Windows 95/NT
SOFTWARE: Stuffit Deluxe 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/087,762
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lavin Jr., Lawrence M.
REGISTRATION NUMBER: 30,768
REFERENCE/DOCKET NUMBER: 04983,0037/38-21(15405)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 737-6670
TELEFAX: (314) 737-6047
INFORMATION FOR SEQ ID NO: 211:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Zea mays
DEVELOPMENTAL STAGE: etiolated 4 days
TISSUE TYPE: seedling
IMMEDIATE SOURCE:
LIBRARY: SATMON029
CLONE: 700454335
US-60-087-762-211

Query Match 3.0%; Score 18; DB 40; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 524 ccacagccgagcggaac 541
|||||
DB 102 CCACAGCCGCGGCGAAC 85

RESULT 50
US-60-092-036-941/C
Sequence 941, Application US/60092036
GENERAL INFORMATION:
APPLICANT: Mitsky, Timothy A.
APPLICANT: Liu, Jingdong
TITLE OF INVENTION: Nucleic acid molecules and other molecules
TITLE OF INVENTION: associated with the Phenylpropanoid Biosynthetic pathway
NUMBER OF SEQUENCES: 2020
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lawrence M. Lavin, Jr.
STREET: BB4F
STREET: Monsanto Company
STREET: 700 Chesterfield Parkway North
CITY: St. Louis
STATE: MO
COUNTRY: United States
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette (3.5in, 1.44 Mb)
COMPUTER: IBM PC/XT/AT, IBM PS/2 or compatibles
OPERATING SYSTEM: Windows 95/NT
SOFTWARE: Winzip 6.3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/092,036
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lavin Jr., Lawrence M.
REGISTRATION NUMBER: 30,768
REFERENCE/DOCKET NUMBER: 04983,0045/38-21(15427)A
TELECOMMUNICATION INFORMATION:

TELEPHONE: (314) 737-6670
 TELEFAX: (314) 737-6047
 INFORMATION FOR SEQ ID NO: 941:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 271 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Zea mays
 DEVELOPMENTAL STAGE: etiolated 4 days
 TISSUE TYPE: seedling
 IMMEDIATE SOURCE:
 LIBRARY: SATMON029
 CLONE: 700454335
 US-60-092-036-941

Query Match 3.0%; Score 18; DB 41; Length 271;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 524 ccacagccgcgcgcgaac 541
 ||||||||||||||||
 DB 102 CCACAGCCGCGCGCAAC 85

Search completed: September 21, 2001, 16:15:12
 Job time: 9257 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 21:10:19 ; Search time 15806.4 Seconds
(without alignments)
955.089 Million cell updates/sec

Title: US-09-138-735-1_COPY_1232_2207
Perfect score: 976

Sequence: 1 caggtacagcgtacagcgtt.....tgtgaacgagccattatgt 976

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1344157 seqs, 7733874588 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl:*

1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_on:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_on:*
43: em_or:*

Oligo search
5+

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 976 | 100.0 | 3402 | 9 | AR047920 |
| 2 | 874 | 89.5 | 3402 | 9 | A48910 |
| 3 | 874 | 89.5 | 3402 | 96 | TCU24190 |
| 4 | 22 | 2.3 | 22 | 9 | A48918 |
| 5 | 22 | 2.3 | 22 | 9 | AR047927 |
| 6 | 22 | 2.3 | 331801 | 3 | NMA422491 |
| 7 | 21 | 2.2 | 21 | 9 | A48919 |
| 8 | 21 | 2.2 | 21 | 9 | A48921 |

| | | | | | | | | | | | | | | | |
|---|----|----|-----|--------|----|-----------|--------------------|---|-----|----|-----|--------|----|--------------|----------------------|
| C | 9 | 21 | 2.2 | 21 | 9 | AR047928 | Sequence | C | 82 | 18 | 1.8 | 58733 | 2 | AF217189 | AF217189 Sorangium |
| C | 10 | 21 | 2.2 | 21 | 9 | AR047930 | Sequence | C | 83 | 18 | 1.8 | 62271 | 13 | ATAC013454 | AC013454 Arabidops |
| C | 11 | 20 | 2.0 | 72056 | 62 | AC011183 | Homo sapi | C | 84 | 18 | 1.8 | 68750 | 2 | AF210843 | AF210843 Sorangium |
| C | 12 | 20 | 2.0 | 91470 | 12 | AC011438 | Genomic s | C | 85 | 18 | 1.8 | 78596 | 12 | AB010695 | AB010695 Arabidops |
| C | 13 | 20 | 2.0 | 115243 | 90 | AL157772 | Human DNA | C | 86 | 18 | 1.8 | 80535 | 87 | AC011992 | AC011992 Arabidops |
| C | 14 | 20 | 2.0 | 153259 | 88 | AC087859 | Homo sapi | C | 87 | 18 | 1.8 | 87045 | 13 | ATAC012393 | ATAC012393 Arabidops |
| C | 15 | 20 | 2.0 | 155075 | 65 | AC018355 | Homo sapi | C | 88 | 18 | 1.8 | 91470 | 12 | AC011438 | AC011438 Genomic s |
| C | 16 | 20 | 2.0 | 171714 | 88 | AC034187 | Homo sapi | C | 89 | 18 | 1.8 | 96424 | 14 | ATP22211 | AL162971 Arabidops |
| C | 17 | 20 | 2.0 | 177816 | 65 | AC017643 | Genomic s | C | 90 | 18 | 1.8 | 97749 | 85 | AC004906 | AC004906 Homo sapi |
| C | 18 | 20 | 2.0 | 182365 | 81 | AL152648 | Genomic s | C | 91 | 18 | 1.8 | 110000 | 61 | AC009579 | Continuation (2 of |
| C | 19 | 20 | 2.0 | 194897 | 4 | AC007697 | Drosophila | C | 92 | 18 | 1.8 | 110000 | 75 | AC073763 | Continuation (2 of |
| C | 20 | 20 | 2.0 | 203195 | 67 | AC022397 | Mus muscu | C | 93 | 18 | 1.8 | 110000 | 84 | LMFLCHR32.03 | Continuation (4 of |
| C | 21 | 20 | 2.0 | 223589 | 69 | AC025210 | Homo sapi | C | 94 | 18 | 1.8 | 110684 | 12 | AC006601 | AC006601 Arabidops |
| C | 22 | 20 | 2.0 | 262395 | 5 | AE003802 | Genomic s | C | 95 | 18 | 1.8 | 113345 | 88 | AF109076 | AF109076 Homo sapi |
| C | 23 | 19 | 1.9 | 1152 | 14 | D89210 | Schizosacch | C | 96 | 18 | 1.8 | 115429 | 81 | AL500525 | AL500525 Homo sapi |
| C | 24 | 19 | 1.9 | 1279 | 94 | AF273691 | Genomic s | C | 97 | 18 | 1.8 | 115925 | 75 | AC074128 | AC074128 Homo sapi |
| C | 25 | 19 | 1.9 | 10128 | 1 | AE005021 | Genomic s | C | 98 | 18 | 1.8 | 124510 | 88 | AC027306 | AC027306 Homo sapi |
| C | 26 | 19 | 1.9 | 11873 | 9 | AR083116 | Sequence | C | 99 | 18 | 1.8 | 127593 | 86 | AC006236 | AC006236 Homo sapi |
| C | 27 | 19 | 1.9 | 11878 | 9 | AR083115 | Sequence | C | 100 | 18 | 1.8 | 129606 | 83 | AP003379 | AP003379 Oryza sat |
| C | 28 | 19 | 1.9 | 11883 | 9 | AR083114 | Sequence | C | 101 | 18 | 1.8 | 132910 | 64 | AC016342 | AC016342 Homo sapi |
| C | 29 | 19 | 1.9 | 67442 | 15 | SPBC262 | Genomic s | C | 102 | 18 | 1.8 | 135303 | 73 | AC068257 | AC068257 Homo sapi |
| C | 30 | 19 | 1.9 | 79970 | 62 | AC011428 | Homo sapi | C | 103 | 18 | 1.8 | 137506 | 85 | AC004824 | AC004824 Homo sapi |
| C | 31 | 19 | 1.9 | 82033 | 94 | AC068803 | Mus muscu | C | 104 | 18 | 1.8 | 139043 | 78 | AC090974 | AC090974 Oryza sat |
| C | 32 | 19 | 1.9 | 95232 | 94 | AC068952 | Mus muscu | C | 105 | 18 | 1.8 | 144057 | 78 | AL138798 | AL138798 Homo sapi |
| C | 33 | 19 | 1.9 | 11000 | 79 | AL354832 | Continuation (2 of | C | 106 | 18 | 1.8 | 147042 | 62 | AC011647 | AC011647 Homo sapi |
| C | 34 | 19 | 1.9 | 11071 | 68 | AC024591 | Homo sapi | C | 107 | 18 | 1.8 | 154461 | 69 | AC025711 | AC025711 Homo sapi |
| C | 35 | 19 | 1.9 | 128562 | 66 | AC020962 | Mus muscu | C | 108 | 18 | 1.8 | 154788 | 66 | AC021498 | AC021498 Homo sapi |
| C | 36 | 19 | 1.9 | 130734 | 94 | AC073882 | Mus muscu | C | 109 | 18 | 1.8 | 157349 | 78 | AF145206 | AF145206 Homo sapi |
| C | 37 | 19 | 1.9 | 131888 | 93 | HS11057D4 | Human DNA | C | 110 | 18 | 1.8 | 157981 | 85 | AC005184 | AC005184 Homo sapi |
| C | 38 | 19 | 1.9 | 134506 | 86 | AC005738 | Homo sapi | C | 111 | 18 | 1.8 | 158427 | 79 | AL355516 | AL355516 Homo sapi |
| C | 39 | 19 | 1.9 | 141158 | 68 | AC023898 | Mus muscu | C | 112 | 18 | 1.8 | 162434 | 79 | AL158169 | AL158169 Homo sapi |
| C | 40 | 19 | 1.9 | 147009 | 76 | AC084179 | Homo sapi | C | 113 | 18 | 1.8 | 164679 | 86 | AC005921 | AC005921 Homo sapi |
| C | 41 | 19 | 1.9 | 156599 | 76 | AC083840 | Homo sapi | C | 114 | 18 | 1.8 | 169195 | 67 | AC023165 | AC023165 Homo sapi |
| C | 42 | 19 | 1.9 | 160404 | 71 | AC034236 | Homo sapi | C | 115 | 18 | 1.8 | 169804 | 71 | AC027741 | AC027741 Homo sapi |
| C | 43 | 19 | 1.9 | 161549 | 64 | AC015815 | Homo sapi | C | 116 | 18 | 1.8 | 173081 | 78 | AL137001 | AL137001 Homo sapi |
| C | 44 | 19 | 1.9 | 170431 | 61 | AC009108 | Homo sapi | C | 117 | 18 | 1.8 | 173843 | 77 | AC084405 | AC084405 Oryza sat |
| C | 45 | 19 | 1.9 | 175249 | 63 | AC013543 | Homo sapi | C | 118 | 18 | 1.8 | 176816 | 62 | AC011372 | AC011372 Homo sapi |
| C | 46 | 19 | 1.9 | 178751 | 80 | AL356772 | Homo sapi | C | 119 | 18 | 1.8 | 177424 | 61 | AC009832 | AC009832 Homo sapi |
| C | 47 | 19 | 1.9 | 179844 | 83 | AP002501 | Homo sapi | C | 120 | 18 | 1.8 | 180788 | 69 | AC025680 | AC025680 Homo sapi |
| C | 48 | 19 | 1.9 | 180717 | 73 | AC068169 | Homo sapi | C | 121 | 18 | 1.8 | 180788 | 76 | AC083757 | AC083757 Homo sapi |
| C | 49 | 19 | 1.9 | 181853 | 68 | AC024507 | Homo sapi | C | 122 | 18 | 1.8 | 180863 | 76 | AC079809 | AC079809 Homo sapi |
| C | 50 | 19 | 1.9 | 184012 | 87 | AC017080 | Homo sapi | C | 123 | 18 | 1.8 | 185437 | 66 | AC020714 | AC020714 Homo sapi |
| C | 51 | 19 | 1.9 | 185380 | 78 | AL136119 | Homo sapi | C | 124 | 18 | 1.8 | 185437 | 77 | AC084356 | AC084356 Homo sapi |
| C | 52 | 19 | 1.9 | 185892 | 78 | AC090652 | Mus muscu | C | 125 | 18 | 1.8 | 185531 | 62 | AC012059 | AC012059 Homo sapi |
| C | 53 | 19 | 1.9 | 191257 | 74 | AC069560 | Mus muscu | C | 126 | 18 | 1.8 | 193444 | 13 | AF229199 | AF229199 Oryza sat |
| C | 54 | 19 | 1.9 | 194575 | 81 | AL512653 | Homo sapi | C | 127 | 18 | 1.8 | 197772 | 60 | AC081129 | AC081129 Homo sapi |
| C | 55 | 19 | 1.9 | 198453 | 62 | AC011359 | Homo sapi | C | 128 | 18 | 1.8 | 199588 | 69 | AC026083 | AC026083 Homo sapi |
| C | 56 | 19 | 1.9 | 200860 | 81 | AL445211 | Homo sapi | C | 129 | 18 | 1.8 | 203200 | 87 | AC008744 | AC008744 Homo sapi |
| C | 57 | 19 | 1.9 | 235141 | 85 | AC004615 | Homo sapi | C | 130 | 18 | 1.8 | 207221 | 62 | AC012564 | AC012564 Homo sapi |
| C | 58 | 19 | 1.9 | 296050 | 79 | AL354813 | Homo sapi | C | 131 | 18 | 1.8 | 217730 | 77 | AC084823 | AC084823 Homo sapi |
| C | 59 | 19 | 1.8 | 18 | 9 | A48914 | Sequence 5 | C | 132 | 18 | 1.8 | 218553 | 76 | AC079940 | AC079940 Mus muscu |
| C | 60 | 18 | 1.8 | 18 | 9 | A48917 | Sequence 8 | C | 133 | 18 | 1.8 | 220469 | 75 | AC010538 | AC010538 Homo sapi |
| C | 61 | 18 | 1.8 | 18 | 9 | AR047923 | Sequence | C | 134 | 18 | 1.8 | 224376 | 61 | AC073714 | AC073714 Mus muscu |
| C | 62 | 18 | 1.8 | 18 | 9 | AR047926 | Sequence | C | 135 | 18 | 1.8 | 230760 | 74 | AC079635 | AC079635 Homo sapi |
| C | 63 | 18 | 1.8 | 18 | 9 | AR047933 | Sequence | C | 136 | 18 | 1.8 | 326419 | 76 | AC079635 | AC079635 Homo sapi |
| C | 64 | 18 | 1.8 | 339 | 53 | AU049339 | Rattus no | C | 137 | 17 | 1.7 | 198 | 58 | AF065721 | AF065721 Stealch v |
| C | 65 | 18 | 1.8 | 939 | 12 | AF134126 | Arabidops | C | 138 | 17 | 1.7 | 339 | 93 | HSNOS2E17 | X85773 H. sapiens N |
| C | 66 | 18 | 1.8 | 942 | 13 | AF143691 | Arabidops | C | 139 | 17 | 1.7 | 368 | 54 | G02521 | G02521 human STS W |
| C | 67 | 18 | 1.8 | 952 | 13 | AF361858 | Arabidops | C | 140 | 17 | 1.7 | 376 | 53 | AU026316 | AU026316 human STS W |
| C | 68 | 18 | 1.8 | 1589 | 10 | AX086831 | Sequence | C | 141 | 17 | 1.7 | 383 | 76 | AC080428 | AC080428 Rattus no |
| C | 69 | 18 | 1.8 | 1589 | 93 | HSW801826 | Homo sapi | C | 142 | 17 | 1.7 | 398 | 54 | G47871 | G47871 Z26401.1 Ze |
| C | 70 | 18 | 1.8 | 2520 | 9 | AR051460 | Sequence | C | 143 | 17 | 1.7 | 400 | 58 | AF174900 | AF174900 HIV-1 Iso |
| C | 71 | 18 | 1.8 | 2520 | 9 | AR072620 | Sequence | C | 144 | 17 | 1.7 | 400 | 58 | AF174963 | AF174963 HIV-1 Iso |
| C | 72 | 18 | 1.8 | 2520 | 9 | AR073165 | Sequence | C | 145 | 17 | 1.7 | 424 | 54 | G24099 | G24099 human STS W |
| C | 73 | 18 | 1.8 | 2723 | 89 | AK026431 | Homo sapi | C | 146 | 17 | 1.7 | 661 | 8 | AF243428 | AF243428 Oncohyne |
| C | 74 | 18 | 1.8 | 4126 | 4 | AC006817 | Genomic s | C | 147 | 17 | 1.7 | 663 | 88 | AF001597 | AF001597 Homo sapi |
| C | 75 | 18 | 1.8 | 9450 | 59 | HPCJ049E1 | Hepatitis C | C | 148 | 17 | 1.7 | 687 | 9 | AX017873 | AX017873 Sequence |
| C | 76 | 18 | 1.8 | 10863 | 1 | AE005678 | Cauliodact | C | 149 | 17 | 1.7 | 804 | 88 | AF110775 | AF110775 Homo sapi |
| C | 77 | 18 | 1.8 | 11031 | 1 | AE004648 | Pseudomon | C | 150 | 17 | 1.7 | 839 | 93 | HS250393 | HS250393 Homo sapi |
| C | 78 | 18 | 1.8 | 14704 | 1 | AE001970 | Deinococc | C | 151 | 17 | 1.7 | 915 | 77 | AC085350 | AC085350 Giardia i |
| C | 79 | 18 | 1.8 | 33800 | 6 | CELC3443 | Caenorhab | C | 152 | 17 | 1.7 | 920 | 71 | AC034857 | AC034857 Giardia i |
| C | 80 | 18 | 1.8 | 38000 | 61 | AC010079 | Leishmani | C | 153 | 17 | 1.7 | 925 | 53 | CNS074H6 | AL428860 Clone BA0 |
| C | 81 | 18 | 1.8 | 42000 | 77 | AC087836 | Leishmani | C | 154 | 17 | 1.7 | 945 | 7 | BOVCCOSD | L08798 Bos taurus |

| | | | | | | | | | | | | | |
|-------|----|-----|-------|----|------------|----------------------|-------|----|-----|--------|----|--------------|------------------------|
| 155 | 17 | 1.7 | 969 | 93 | MSM800753 | AL080224 Homo sapi | c 228 | 17 | 1.7 | 35525 | 77 | AC087068 | AC087068 Leishmani |
| 156 | 17 | 1.7 | 977 | 94 | MMU250394 | AZ250394 Mus muscu | 229 | 17 | 1.7 | 35896 | 12 | AB024027 | AB024027 Arabidops |
| 157 | 17 | 1.7 | 990 | 53 | CNS0722C | AT426862 Clone BA0 | 230 | 17 | 1.7 | 36849 | 3 | SC1C3 | AL023702 Streptomy |
| 158 | 17 | 1.7 | 993 | 5 | AP190457 | AP190457 Trypanoso | c 231 | 17 | 1.7 | 37056 | 85 | AC004700 | AC004700 Homo sapi |
| c 159 | 17 | 1.7 | 1009 | 71 | AC029282 | AC029282 Giardia i | c 232 | 17 | 1.7 | 38024 | 6 | CELF52H1 | U41990 Caenorhabdi |
| c 160 | 17 | 1.7 | 1046 | 89 | AF161497 | AF161497 Homo sapi | c 233 | 17 | 1.7 | 38919 | 86 | AC005947 | AC005947 Homo sapi |
| c 161 | 17 | 1.7 | 1067 | 53 | CNS07BiH | AL437919 T7 end of | c 234 | 17 | 1.7 | 39207 | 3 | SC4G1 | AL331039 Streptomy |
| c 162 | 17 | 1.7 | 1062 | 94 | BC004726 | BC004726 Mus muscu | c 235 | 17 | 1.7 | 39266 | 86 | AC005791 | AC005791 Homo sapi |
| 163 | 17 | 1.7 | 1063 | 5 | AF316150 | AF316150 Trypanoso | c 236 | 17 | 1.7 | 39436 | 85 | AB000879 | AB000879 Homo sapi |
| 164 | 17 | 1.7 | 1151 | 53 | CNS06TMT | AL414747 T3 end of | c 237 | 17 | 1.7 | 40121 | 3 | MSGF414A | AD000007 Mycobacte |
| c 165 | 17 | 1.7 | 1181 | 97 | HUMMHI | L93376 Homo sapien | c 238 | 17 | 1.7 | 40673 | 6 | LMFL377 | AL136332 Leishmani |
| 166 | 17 | 1.7 | 1449 | 1 | AF075168 | AF075168 Brucella | c 239 | 17 | 1.7 | 40881 | 85 | AC004679 | AC004679 Homo sapi |
| c 167 | 17 | 1.7 | 1454 | 2 | BSP29998 | Z99998 Uncultured | c 240 | 17 | 1.7 | 40910 | 85 | AC005260 | AC005260 Homo sapi |
| c 168 | 17 | 1.7 | 1497 | 8 | AF247363 | AF247363 Paleosuch | c 241 | 17 | 1.7 | 42139 | 85 | AC004654 | AC004654 Homo sapi |
| c 169 | 17 | 1.7 | 1516 | 15 | ZMA5702 | Y13733 Zee mays mr | c 242 | 17 | 1.7 | 42467 | 6 | CELF26E12 | U55373 Caenorhabdi |
| c 170 | 17 | 1.7 | 1517 | 15 | AF309358 | AF309358 Hordelum v | c 243 | 17 | 1.7 | 42584 | 89 | AF215839 | AF215839 Homo sapi |
| c 171 | 17 | 1.7 | 1671 | 53 | RS06694 | U303632 Chlamydomon | c 244 | 17 | 1.7 | 42649 | 88 | AC024077 | AC024077 Homo sapi |
| c 172 | 17 | 1.7 | 1675 | 14 | CRU03632 | U303632 S. macrospor | c 245 | 17 | 1.7 | 43147 | 3 | SC4A10 | AL109663 Streptomy |
| c 173 | 17 | 1.7 | 2242 | 15 | SMTUBA | U63336 Human MHC C | c 246 | 17 | 1.7 | 43401 | 3 | MTY13E12 | Z95390 Mycobacteri |
| c 174 | 17 | 1.7 | 2261 | 97 | HSU63336 | AL117521 Homo sapi | c 247 | 17 | 1.7 | 43430 | 3 | MTY13E12 | Z81331 Mycobacteri |
| c 175 | 17 | 1.7 | 2271 | 93 | MSM801043 | AL117521 Homo sapi | c 248 | 17 | 1.7 | 43600 | 85 | AC004209 | AC004209 Homo sapi |
| c 176 | 17 | 1.7 | 2298 | 94 | AF177346 | AF177346 Mus muscu | c 249 | 17 | 1.7 | 43690 | 85 | AC005238 | AC005238 Homo sapi |
| c 177 | 17 | 1.7 | 2341 | 89 | AF119857 | AF119857 Homo sapi | c 250 | 17 | 1.7 | 43784 | 64 | AC017442 | AC017442 Drosophill |
| c 178 | 17 | 1.7 | 2402 | 89 | AK022628 | AK022628 Homo sapi | c 251 | 17 | 1.7 | 44710 | 85 | AC004184 | AC004184 Homo sapi |
| c 179 | 17 | 1.7 | 2459 | 10 | AX077260 | AX077260 Sequence | c 252 | 17 | 1.7 | 44760 | 85 | AC004784 | AC004784 Homo sapi |
| c 180 | 17 | 1.7 | 2474 | 94 | MMU278891 | A3278891 Mus muscu | c 253 | 17 | 1.7 | 45087 | 82 | AL590645 | AL590645 Homo sapi |
| c 181 | 17 | 1.7 | 2545 | 9 | AR035063 | AR035063 Sequence | c 254 | 17 | 1.7 | 45344 | 90 | AL356497 | AL356497 Human DNA |
| c 182 | 17 | 1.7 | 2545 | 93 | SHHMDIG | X72755 H. sapiens H | c 255 | 17 | 1.7 | 51082 | 78 | AC090454 | AC090454 Homo sapi |
| c 183 | 17 | 1.7 | 2545 | 58 | AF247522 | AF247522 HIV-1 Iso | c 256 | 17 | 1.7 | 57399 | 91 | AP000433 | AP000433 Homo sapi |
| c 184 | 17 | 1.7 | 2833 | 6 | DR00DIF | L29015 Drosophilla | c 257 | 17 | 1.7 | 57597 | 77 | AC090035 | AC090035 Homo sapi |
| c 185 | 17 | 1.7 | 2920 | 91 | BC002350 | BC002350 Homo sapi | c 258 | 17 | 1.7 | 65174 | 69 | AC025214 | AC025214 Homo sapi |
| c 186 | 17 | 1.7 | 2928 | 89 | AK025328 | AK025328 Homo sapi | c 259 | 17 | 1.7 | 66066 | 70 | AC026566 | AC026566 Homo sapi |
| c 187 | 17 | 1.7 | 2984 | 91 | BC002482 | BC002482 Homo sapi | c 260 | 17 | 1.7 | 66688 | 77 | AC087757 | AC087757 Homo sapi |
| c 188 | 17 | 1.7 | 3240 | 3 | LHPEP1GN | Z56283 L. helveticu | c 261 | 17 | 1.7 | 67392 | 65 | AC018284 | AC018284 Drosophill |
| c 189 | 17 | 1.7 | 3337 | 10 | I23337 | I23337 Sequence 1 | c 262 | 17 | 1.7 | 68723 | 75 | AC079066 | AC079066 Homo sapi |
| c 190 | 17 | 1.7 | 3348 | 2 | D83672 | D83672 Streptomyce | c 263 | 17 | 1.7 | 74371 | 85 | AC005369 | AC005369 Homo sapi |
| c 191 | 17 | 1.7 | 3453 | 3 | MTU43540 | U43540 Mycobacteri | c 264 | 17 | 1.7 | 80099 | 13 | AP002040 | AP002040 Arabidops |
| c 192 | 17 | 1.7 | 3453 | 3 | SSU38802 | U38802 Synecocyst | c 265 | 17 | 1.7 | 82604 | 91 | AL500522 | AL500522 Homo sapi |
| c 193 | 17 | 1.7 | 4308 | 3 | AB056760 | AB056760 Macaca fa | c 266 | 17 | 1.7 | 84957 | 12 | AC006620 | AC006620 Arabidops |
| c 194 | 17 | 1.7 | 4501 | 85 | AB030448S1 | AB030448 Mus muscu | c 267 | 17 | 1.7 | 85657 | 12 | AC069473 | AC069473 Arabidops |
| c 195 | 17 | 1.7 | 4545 | 94 | PAU67315 | U67315 Pseudomonas | c 268 | 17 | 1.7 | 86155 | 90 | AL159169 | AL159169 Human DNA |
| c 196 | 17 | 1.7 | 4601 | 3 | AF064104 | AF064104 Homo sapi | c 269 | 17 | 1.7 | 88296 | 88 | AC027328 | AC027328 Homo sapi |
| c 197 | 17 | 1.7 | 4824 | 88 | AK001661 | AK001661 Homo sapi | c 270 | 17 | 1.7 | 88401 | 14 | F23M19 | AC007454 Arabidops |
| c 198 | 17 | 1.7 | 4823 | 89 | AF064105 | AF064105 Homo sapi | c 271 | 17 | 1.7 | 88401 | 70 | AC026865 | AC026865 Homo sapi |
| c 199 | 17 | 1.7 | 5186 | 3 | SADNAIE2 | X87940 S. aurantiac | c 272 | 17 | 1.7 | 88784 | 63 | AC012997 | AC012997 Drosophill |
| c 200 | 17 | 1.7 | 5851 | 94 | AF15253S1 | AF152533 Mus muscu | c 273 | 17 | 1.7 | 90767 | 78 | H0510X06 | AF257419 Oryza sat |
| c 201 | 17 | 1.7 | 6074 | 94 | MUSP3VPR1 | AF098866 Mus sp. 1 | c 274 | 17 | 1.7 | 94641 | 84 | AF257499 | AF257499 Homo sapi |
| c 202 | 17 | 1.7 | 6167 | 85 | AB023148 | AB023148 Homo sapi | c 275 | 17 | 1.7 | 98758 | 92 | HS223H9 | AL136081 Human DNA |
| c 203 | 17 | 1.7 | 6667 | 7 | BOYCOX7AL | L09603 Bovine cyto | c 276 | 17 | 1.7 | 99850 | 89 | AL136081 | AL136081 Human DNA |
| c 204 | 17 | 1.7 | 8657 | 8 | GGU35037 | U35037 Gallus gall | c 277 | 17 | 1.7 | 100000 | 91 | AP000506 | AP000506 Homo sapi |
| c 205 | 17 | 1.7 | 9083 | 58 | AF193253 | AF193253 HIV-1 sub | c 278 | 17 | 1.7 | 100000 | 91 | AP000513 | AP000513 Homo sapi |
| c 206 | 17 | 1.7 | 10599 | 1 | AE000939 | AE000939 Methanoba | c 279 | 17 | 1.7 | 103419 | 78 | AL139221 | AL139221 Homo sapi |
| c 207 | 17 | 1.7 | 10599 | 1 | AE001037 | AE001037 Archaeogl | c 280 | 17 | 1.7 | 104597 | 92 | HS393P12 | HS393P12 Homo sapi |
| c 208 | 17 | 1.7 | 11008 | 1 | AE004505 | AE004505 Pseudomon | c 281 | 17 | 1.7 | 107150 | 60 | AC008839 | AC008839 Homo sapi |
| c 209 | 17 | 1.7 | 11421 | 2 | ECOPHN | D90227 Escherichia | c 282 | 17 | 1.7 | 107152 | 88 | AC020930 | AC020930 Homo sapi |
| c 210 | 17 | 1.7 | 11672 | 2 | ECOPHN | AE005822 Caulobact | c 283 | 17 | 1.7 | 109279 | 92 | HS394F12 | HS394F12 Homo sapi |
| c 211 | 17 | 1.7 | 11908 | 1 | AE005822 | AE004441 Pseudomon | c 284 | 17 | 1.7 | 109626 | 87 | AC008727 | AC008727 Homo sapi |
| c 212 | 17 | 1.7 | 12213 | 1 | AE004441 | AE005375 Escherich | c 285 | 17 | 1.7 | 109626 | 87 | AC010346 | AC010346 Homo sapi |
| c 213 | 17 | 1.7 | 12857 | 1 | AE005375 | AE004049 Xylella f | c 286 | 17 | 1.7 | 110000 | 83 | CEY81G3_0 | CEY81G3_0 Homo sapi |
| c 214 | 17 | 1.7 | 15552 | 1 | AE004049 | J05260 E. coli psid | c 287 | 17 | 1.7 | 110000 | 84 | LMFLCHR16_03 | LMFLCHR16_03 Homo sapi |
| c 215 | 17 | 1.7 | 15611 | 2 | ECOPHNAQ | AL035212 Streptomy | c 288 | 17 | 1.7 | 110000 | 84 | LMFLCHR25_03 | LMFLCHR25_03 Homo sapi |
| c 216 | 17 | 1.7 | 20102 | 6 | CEY81G3A | AL035453 Caenorhab | c 289 | 17 | 1.7 | 110000 | 84 | LMFLCHR36_09 | LMFLCHR36_09 Homo sapi |
| c 217 | 17 | 1.7 | 20303 | 6 | CEY81G3A | AE000482 Escherich | c 290 | 17 | 1.7 | 110000 | 84 | LMFLCHR36_11 | LMFLCHR36_11 Homo sapi |
| c 218 | 17 | 1.7 | 20906 | 1 | AE000482 | AL163672 Streptomy | c 291 | 17 | 1.7 | 110000 | 84 | LMFLCHR36_12 | LMFLCHR36_12 Homo sapi |
| c 219 | 17 | 1.7 | 27887 | 4 | AE002700 | AE002700 Drosophill | c 292 | 17 | 1.7 | 110603 | 93 | HSJ421D16 | HSJ421D16 Homo sapi |
| c 220 | 17 | 1.7 | 27887 | 4 | AE002700 | AB023052 Homo sapi | c 293 | 17 | 1.7 | 117273 | 63 | AC010662 | AC010662 Drosophill |
| c 221 | 17 | 1.7 | 28139 | 65 | AC020057 | AC020057 Drosophill | c 294 | 17 | 1.7 | 117273 | 63 | AC015138 | AC015138 Homo sapi |
| c 222 | 17 | 1.7 | 29139 | 65 | AC020057 | Z92770 Mycobacteri | c 295 | 17 | 1.7 | 119919 | 89 | AF241734 | AF241734 Drosophill |
| c 223 | 17 | 1.7 | 30200 | 15 | SPAC2068 | Z95334 S. pombe chr | c 296 | 17 | 1.7 | 122258 | 80 | AL356427 | AL356427 Homo sapi |
| c 224 | 17 | 1.7 | 31015 | 3 | MTY15 | Z92770 Mycobacteri | c 297 | 17 | 1.7 | 122258 | 80 | AL356427 | AL356427 Homo sapi |
| c 225 | 17 | 1.7 | 33451 | 12 | CEYK20 | Z69904 Caenorhabdi | c 298 | 17 | 1.7 | 123085 | 90 | AL157406 | AL157406 Human DNA |
| c 226 | 17 | 1.7 | 33490 | 6 | CEYK20 | Z69904 Caenorhabdi | c 299 | 17 | 1.7 | 126152 | 92 | HS433M19 | HS433M19 Homo sapi |
| c 227 | 17 | 1.7 | 33758 | 6 | CEYK20 | Z50806 Caenorhabdi | c 300 | 17 | 1.7 | 132275 | 72 | AC048388 | AC048388 Homo sapi |

| | | | | | | | | | | | |
|-------|----|------------|----|-----------|---------------------|-------|----|------------|----|-----------|---------------------|
| C 301 | 17 | 1.7 133751 | 92 | HS1131C10 | AL121931 Human DNA | 374 | 17 | 1.7 167237 | 66 | AC021609 | AC021609 Homo sapi |
| C 302 | 17 | 1.7 137570 | 77 | AC084844 | AC084844 Homo sapi | 375 | 17 | 1.7 167310 | 73 | AC068972 | AC068972 Homo sapi |
| C 303 | 17 | 1.7 137928 | 78 | AC090419 | AC090419 Homo sapi | 376 | 17 | 1.7 167555 | 64 | AC016044 | AC016044 Homo sapi |
| C 304 | 17 | 1.7 137943 | 76 | AC084032 | Oryza sat | C 377 | 17 | 1.7 167590 | 80 | AL356221 | AL356221 Homo sapi |
| C 305 | 17 | 1.7 138141 | 65 | AC018433 | AC018433 Homo sapi | C 378 | 17 | 1.7 167590 | 82 | AC046148 | AC046148 Homo sapi |
| C 306 | 17 | 1.7 138675 | 92 | HS1057B20 | AL109822 Human DNA | C 379 | 17 | 1.7 169241 | 81 | AL445987 | AL445987 Homo sapi |
| C 307 | 17 | 1.7 139505 | 93 | HS0591C20 | AL118506 Human DNA | C 380 | 17 | 1.7 169256 | 76 | AC048358 | AC048358 Homo sapi |
| C 308 | 17 | 1.7 140154 | 72 | AC055843 | AL118506 Homo sapi | C 381 | 17 | 1.7 169256 | 82 | AC005829 | AC005829 Homo sapi |
| C 309 | 17 | 1.7 141017 | 12 | AC037197 | AC037197 Oryza sat | C 382 | 17 | 1.7 170324 | 4 | AC007805 | AC007805 Drosophill |
| C 310 | 17 | 1.7 141924 | 89 | AL138830 | AL138830 Human DNA | C 383 | 17 | 1.7 170324 | 4 | AC007805 | AC007805 Drosophill |
| C 311 | 17 | 1.7 142203 | 76 | AC079521 | AC079521 Mus muscu | C 384 | 17 | 1.7 170431 | 61 | AC009108 | AC009108 Homo sapi |
| C 312 | 17 | 1.7 143759 | 88 | AC073532 | AC073532 Homo sapi | C 385 | 17 | 1.7 170457 | 76 | AC083944 | AC083944 Oryza sat |
| C 313 | 17 | 1.7 144316 | 79 | AL162582 | AL162592 Homo sapi | C 386 | 17 | 1.7 170711 | 71 | AC032020 | AC032020 Homo sapi |
| C 314 | 17 | 1.7 144704 | 64 | AC016595 | AC016595 Homo sapi | C 387 | 17 | 1.7 170854 | 83 | AP002383 | AP002383 Homo sapi |
| C 315 | 17 | 1.7 145642 | 71 | AC027805 | AC027805 Homo sapi | C 388 | 17 | 1.7 170856 | 73 | AC068160 | AC068160 Homo sapi |
| C 316 | 17 | 1.7 146018 | 65 | AC018404 | AC018404 Homo sapi | C 389 | 17 | 1.7 170872 | 69 | AC004951 | AC004951 Homo sapi |
| C 317 | 17 | 1.7 146271 | 3 | SYCS1RB | DE4000 Synchocyst | C 390 | 17 | 1.7 170891 | 66 | AC002118 | AC002118 Homo sapi |
| C 318 | 17 | 1.7 146393 | 70 | AC026791 | AC026791 Homo sapi | C 391 | 17 | 1.7 171539 | 81 | AL512788 | AL512788 Homo sapi |
| C 319 | 17 | 1.7 148960 | 85 | AB000882 | AB000882 Homo sapi | C 392 | 17 | 1.7 171775 | 70 | AC027015 | AC027015 Homo sapi |
| C 320 | 17 | 1.7 149143 | 62 | AC011591 | AC011591 Homo sapi | C 393 | 17 | 1.7 171790 | 62 | AC011106 | AC011106 Homo sapi |
| C 321 | 17 | 1.7 149773 | 71 | AC041029 | AC041029 Homo sapi | C 394 | 17 | 1.7 171917 | 82 | AL590638 | AL590638 Homo sapi |
| C 322 | 17 | 1.7 150250 | 70 | AC027573 | AC027573 Homo sapi | C 395 | 17 | 1.7 172004 | 86 | AC008482 | AC008482 Homo sapi |
| C 323 | 17 | 1.7 150399 | 63 | AC015784 | AC015784 Homo sapi | C 396 | 17 | 1.7 172024 | 75 | AC078971 | AC078971 Homo sapi |
| C 324 | 17 | 1.7 151289 | 83 | AP003228 | Oryza sat | C 397 | 17 | 1.7 172312 | 80 | AL359822 | AL359822 Homo sapi |
| C 325 | 17 | 1.7 151486 | 83 | AP003312 | AP003312 Oryza sat | C 398 | 17 | 1.7 172650 | 62 | AC011042 | AC011042 Homo sapi |
| C 326 | 17 | 1.7 151497 | 71 | AC027809 | AC027809 Homo sapi | C 399 | 17 | 1.7 173235 | 62 | AC023756 | AC023756 Homo sapi |
| C 327 | 17 | 1.7 151654 | 62 | AC012265 | AC012265 Homo sapi | C 400 | 17 | 1.7 173254 | 68 | AC012057 | AC012057 Homo sapi |
| C 328 | 17 | 1.7 151983 | 70 | AC027178 | AC027178 Homo sapi | C 401 | 17 | 1.7 173425 | 65 | AC021482 | AC021482 Homo sapi |
| C 329 | 17 | 1.7 152159 | 70 | AC027334 | AC027334 Homo sapi | C 402 | 17 | 1.7 173973 | 76 | AC073871 | AC073871 Homo sapi |
| C 330 | 17 | 1.7 152513 | 83 | AP003207 | AP003207 Oryza sat | C 403 | 17 | 1.7 174006 | 72 | AC027062 | AC027062 Homo sapi |
| C 331 | 17 | 1.7 152517 | 87 | AC009468 | AC009468 Homo sapi | C 404 | 17 | 1.7 174250 | 70 | AC063950 | AC063950 Homo sapi |
| C 332 | 17 | 1.7 152714 | 80 | AL357042 | AL357042 Homo sapi | C 405 | 17 | 1.7 174476 | 75 | AC079171 | AC079171 Homo sapi |
| C 333 | 17 | 1.7 154254 | 66 | AC021360 | AC021360 Homo sapi | C 406 | 17 | 1.7 174517 | 72 | AC087831 | AC087831 Homo sapi |
| C 334 | 17 | 1.7 154685 | 86 | AC007214 | AC007214 Pan trogl | C 407 | 17 | 1.7 174831 | 77 | AC060809 | AC060809 Homo sapi |
| C 335 | 17 | 1.7 154745 | 61 | AC010860 | AC010860 Homo sapi | C 408 | 17 | 1.7 174852 | 81 | AC010195 | AC010195 Homo sapi |
| C 336 | 17 | 1.7 155068 | 61 | AC027307 | AC027307 Homo sapi | C 409 | 17 | 1.7 175191 | 87 | AC010251 | AC010251 Homo sapi |
| C 337 | 17 | 1.7 155258 | 84 | HS66C13 | AL354047 Homo sapi | C 410 | 17 | 1.7 175202 | 70 | AC027254 | AC027254 Homo sapi |
| C 338 | 17 | 1.7 155399 | 69 | AC025663 | AC025663 Homo sapi | C 411 | 17 | 1.7 175706 | 63 | AC013798 | AC013798 Homo sapi |
| C 339 | 17 | 1.7 155645 | 66 | AC021153 | AC021153 Homo sapi | C 412 | 17 | 1.7 176300 | 71 | AC027708 | AC027708 Homo sapi |
| C 340 | 17 | 1.7 155813 | 65 | AC019350 | AC019350 Homo sapi | C 413 | 17 | 1.7 176552 | 72 | AC079452 | AC079452 Homo sapi |
| C 341 | 17 | 1.7 156241 | 80 | AL358034 | AL358034 Homo sapi | C 414 | 17 | 1.7 176683 | 72 | AC055801 | AC055801 Homo sapi |
| C 342 | 17 | 1.7 156557 | 63 | AC015545 | AC015545 Homo sapi | C 415 | 17 | 1.7 177364 | 86 | AC005409 | AC005409 Homo sapi |
| C 343 | 17 | 1.7 157189 | 4 | AC008307 | AC008307 Drosophill | C 416 | 17 | 1.7 177654 | 85 | AC025988 | AC025988 Homo sapi |
| C 344 | 17 | 1.7 158219 | 86 | AC005534 | AC005534 Homo sapi | C 417 | 17 | 1.7 177656 | 76 | AC079911 | AC079911 Mus muscu |
| C 345 | 17 | 1.7 158447 | 75 | AC079071 | AC079071 Homo sapi | C 418 | 17 | 1.7 177754 | 64 | AC021034 | AC021034 Homo sapi |
| C 346 | 17 | 1.7 158458 | 86 | AC007395 | AC007395 Homo sapi | C 419 | 17 | 1.7 177826 | 66 | AC016852 | AC016852 Homo sapi |
| C 347 | 17 | 1.7 158646 | 79 | AL354708 | AL354708 Homo sapi | C 420 | 17 | 1.7 177942 | 81 | AL391803 | AL391803 Homo sapi |
| C 348 | 17 | 1.7 158969 | 81 | AL391826 | AL391826 Homo sapi | C 421 | 17 | 1.7 178395 | 81 | AL512664 | AL512664 Homo sapi |
| C 349 | 17 | 1.7 159007 | 74 | AC073059 | AC073059 Homo sapi | C 422 | 17 | 1.7 178518 | 69 | AC025332 | AC025332 Homo sapi |
| C 350 | 17 | 1.7 159452 | 63 | AC015746 | AC015746 Homo sapi | C 423 | 17 | 1.7 178630 | 64 | AC016219 | AC016219 Homo sapi |
| C 351 | 17 | 1.7 160972 | 4 | AC007723 | AC007723 Drosophill | C 424 | 17 | 1.7 179471 | 75 | AC078868 | AC078868 Homo sapi |
| C 352 | 17 | 1.7 161087 | 65 | AC018850 | AC018850 Homo sapi | C 425 | 17 | 1.7 180340 | 65 | AC016233 | AC016233 Homo sapi |
| C 353 | 17 | 1.7 162014 | 75 | AC074323 | AC074323 Homo sapi | C 426 | 17 | 1.7 180530 | 65 | AC018613 | AC018613 Homo sapi |
| C 354 | 17 | 1.7 162066 | 73 | AC068317 | AC068317 Homo sapi | C 427 | 17 | 1.7 182244 | 74 | AC069150 | AC069150 Homo sapi |
| C 355 | 17 | 1.7 162427 | 87 | AC010205 | AC010205 Homo sapi | C 428 | 17 | 1.7 182718 | 64 | AC017095 | AC017095 Homo sapi |
| C 356 | 17 | 1.7 162691 | 94 | AC003061 | AC003061 Mouse Chr | C 429 | 17 | 1.7 183019 | 70 | AC027564 | AC027564 Homo sapi |
| C 357 | 17 | 1.7 162853 | 13 | AP003285 | AP003285 Oryza sat | C 430 | 17 | 1.7 183251 | 82 | AP001355 | AP001355 Homo sapi |
| C 358 | 17 | 1.7 163134 | 65 | AC019286 | AC019286 Homo sapi | C 431 | 17 | 1.7 183615 | 83 | AP001494 | AP001494 Homo sapi |
| C 359 | 17 | 1.7 163570 | 87 | AC016910 | AC016910 Homo sapi | C 432 | 17 | 1.7 184165 | 81 | AL450317 | AL450317 Mus muscu |
| C 360 | 17 | 1.7 163578 | 74 | AC073657 | AC073657 Homo sapi | C 433 | 17 | 1.7 184651 | 80 | AL450317 | AL450317 Mus muscu |
| C 361 | 17 | 1.7 163836 | 88 | AC073917 | AC073917 Homo sapi | C 434 | 17 | 1.7 185260 | 64 | AC016065 | AC016065 Homo sapi |
| C 362 | 17 | 1.7 163861 | 80 | AL356059 | AL356059 Homo sapi | C 435 | 17 | 1.7 185551 | 93 | HS0171K16 | HS0171K16 Human DNA |
| C 363 | 17 | 1.7 163894 | 62 | AC011985 | AC011985 Homo sapi | C 436 | 17 | 1.7 185948 | 67 | AC022370 | AC022370 Homo sapi |
| C 364 | 17 | 1.7 164034 | 87 | AC011489 | AC011489 Homo sapi | C 437 | 17 | 1.7 186102 | 83 | AP001983 | AP001983 Homo sapi |
| C 365 | 17 | 1.7 164408 | 83 | CNS01DRF | AL117190 Homo sapi | C 438 | 17 | 1.7 186208 | 65 | AC019204 | AC019204 Homo sapi |
| C 366 | 17 | 1.7 165405 | 61 | AC009420 | AC009420 Homo sapi | C 439 | 17 | 1.7 187965 | 74 | AC017010 | AC017010 Homo sapi |
| C 367 | 17 | 1.7 165443 | 76 | AC079939 | AC079939 Mus muscu | C 440 | 17 | 1.7 188463 | 64 | AC073413 | AC073413 Homo sapi |
| C 368 | 17 | 1.7 165725 | 68 | AC024380 | AC024380 Homo sapi | C 441 | 17 | 1.7 188526 | 62 | AC012342 | AC012342 Homo sapi |
| C 369 | 17 | 1.7 165764 | 67 | AC022527 | AC022527 Homo sapi | C 442 | 17 | 1.7 188713 | 87 | AC024167 | AC024167 Homo sapi |
| C 370 | 17 | 1.7 165817 | 60 | AC007935 | AC007935 Homo sapi | C 443 | 17 | 1.7 188755 | 87 | AC013562 | AC013562 Homo sapi |
| C 371 | 17 | 1.7 166107 | 68 | AC024427 | AC024427 Homo sapi | C 444 | 17 | 1.7 188812 | 73 | AC013419 | AC013419 Homo sapi |
| C 372 | 17 | 1.7 167020 | 87 | AC010238 | AC010238 Homo sapi | C 445 | 17 | 1.7 189333 | 66 | AC079445 | AC079445 Mus muscu |
| C 373 | 17 | 1.7 167083 | 70 | AC027459 | AC027459 Homo sapi | C 446 | 17 | 1.7 189496 | 77 | AC084353 | AC084353 Homo sapi |

| | | | | | | | | | | | |
|-----|----|------------|----|------------|-------------|-----|----|------------|----|------------|------------|
| 447 | 17 | 1.7 189786 | 73 | AC068015 | Homo sapi | 520 | 17 | 1.7 310360 | 75 | AC074172 | Mus muscu |
| 448 | 17 | 1.7 190386 | 75 | AC074216 | Homo sapi | 521 | 17 | 1.7 310565 | 79 | CEY62E5 | |
| 449 | 17 | 1.7 190844 | 75 | AC073990 | Homo sapi | 522 | 17 | 1.7 311771 | 83 | AL354714 | Homo sapi |
| 450 | 17 | 1.7 191111 | 86 | AC007458 | Homo sapi | 523 | 17 | 1.7 334666 | 80 | AL359078 | Homo sapi |
| 451 | 17 | 1.7 191352 | 74 | AC073565 | Mus muscu | 524 | 17 | 1.7 337101 | 97 | HSKSRPXR | |
| 452 | 17 | 1.7 192431 | 67 | AC022059 | Homo sapi | 525 | 17 | 1.7 338534 | 2 | EC00W93 | |
| 453 | 17 | 1.7 192868 | 88 | AC023484 | Homo sapi | 526 | 17 | 1.7 339485 | 13 | AF172282 | Oryza sat |
| 454 | 17 | 1.7 193098 | 89 | AL137140 | Human DNA | 527 | 17 | 1.7 340000 | 91 | AF001675 | Homo sapi |
| 455 | 17 | 1.7 193426 | 73 | AC067942 | Homo sapi | 528 | 16 | 1.6 411300 | 10 | 114300 | Sequence 1 |
| 456 | 17 | 1.7 193503 | 79 | AL139818 | Homo sapi | 529 | 16 | 1.6 77 | 88 | AF016206 | Homo sapi |
| 457 | 17 | 1.7 193816 | 69 | AC025874 | Mus muscu | 530 | 16 | 1.6 222 | 10 | E08211 | |
| 458 | 17 | 1.7 194151 | 66 | AC021185 | Homo sapi | 531 | 16 | 1.6 222 | 14 | OSTOS11A | |
| 459 | 17 | 1.7 194159 | 75 | AC074177 | Mus muscu | 532 | 16 | 1.6 224 | 14 | ROALPITS2 | |
| 460 | 17 | 1.7 194812 | 76 | AC079858 | Homo sapi | 533 | 16 | 1.6 224 | 14 | ROADRITS2 | |
| 461 | 17 | 1.7 195310 | 68 | AC024189 | Homo sapi | 534 | 16 | 1.6 224 | 14 | ROBRARITS2 | |
| 462 | 17 | 1.7 195418 | 68 | AC023524 | Homo sapi | 535 | 16 | 1.6 224 | 14 | ROGANITS2 | |
| 463 | 17 | 1.7 195466 | 71 | AC032024 | Homo sapi | 536 | 16 | 1.6 224 | 14 | ROPURITS2 | |
| 464 | 17 | 1.7 195472 | 65 | AC018398 | Homo sapi | 537 | 16 | 1.6 224 | 14 | ROTUMITS2 | |
| 465 | 17 | 1.7 196556 | 65 | AC018382 | Homo sapi | 538 | 16 | 1.6 225 | 14 | CGRACITS2 | |
| 466 | 17 | 1.7 196900 | 66 | AC020851 | Mus muscu | 539 | 16 | 1.6 225 | 14 | CSPTCITS2 | |
| 467 | 17 | 1.7 197871 | 73 | AC068053 | Homo sapi | 540 | 16 | 1.6 225 | 14 | ROAUSITS2 | |
| 468 | 17 | 1.7 198388 | 74 | AC073426 | Homo sapi | 541 | 16 | 1.6 225 | 14 | ROCAUITS2 | |
| 469 | 17 | 1.7 198677 | 1 | AE001863 | Delinococc | 542 | 16 | 1.6 225 | 14 | ROHUMITS2 | |
| 470 | 17 | 1.7 199243 | 71 | AC040925 | Homo sapi | 543 | 16 | 1.6 225 | 14 | ROPRARITS2 | |
| 471 | 17 | 1.7 199314 | 61 | AC009097 | Homo sapi | 544 | 16 | 1.6 225 | 14 | ROSCHITS2 | |
| 472 | 17 | 1.7 199853 | 75 | AC073785 | Mus muscu | 545 | 16 | 1.6 225 | 14 | ROSCITTS2 | |
| 473 | 17 | 1.7 200033 | 93 | HSA251973 | Homo sapi | 546 | 16 | 1.6 225 | 14 | ROTIIRITS2 | |
| 474 | 17 | 1.7 200349 | 97 | HSU85195 | Homo sapien | 547 | 16 | 1.6 225 | 14 | ROMARITS2 | |
| 475 | 17 | 1.7 201660 | 69 | AC025105 | Homo sapi | 548 | 16 | 1.6 227 | 53 | ROARITS2 | |
| 476 | 17 | 1.7 201846 | 60 | AC009036 | Homo sapi | 549 | 16 | 1.6 270 | 14 | AY021672 | |
| 477 | 17 | 1.7 205251 | 85 | AC005392 | Homo sapi | 550 | 16 | 1.6 293 | 54 | G04984 | |
| 478 | 17 | 1.7 206518 | 74 | AC069091 | Homo sapi | 551 | 16 | 1.6 299 | 92 | HS98C10R | |
| 479 | 17 | 1.7 207083 | 72 | AC055797 | Homo sapi | 552 | 16 | 1.6 334 | 89 | AX000736 | |
| 480 | 17 | 1.7 207103 | 73 | AC067731 | Homo sapi | 553 | 16 | 1.6 334 | 89 | AF147324 | |
| 481 | 17 | 1.7 207170 | 83 | AP003484 | Homo sapi | 554 | 16 | 1.6 336 | 59 | HV1ETG4 | |
| 482 | 17 | 1.7 207411 | 77 | AC084746 | Mus muscu | 555 | 16 | 1.6 341 | 58 | AF028579 | |
| 483 | 17 | 1.7 209034 | 75 | AC073789 | Mus muscu | 556 | 16 | 1.6 343 | 59 | AF028578 | |
| 484 | 17 | 1.7 209823 | 71 | AC044907 | Homo sapi | 557 | 16 | 1.6 343 | 59 | HV1S31781 | |
| 485 | 17 | 1.7 211959 | 81 | AL390761 | Homo sapi | 558 | 16 | 1.6 343 | 59 | HV1S31781 | |
| 486 | 17 | 1.7 212209 | 68 | AC024651 | Homo sapi | 559 | 16 | 1.6 343 | 59 | HV1S31783 | |
| 487 | 17 | 1.7 217115 | 60 | AC007305 | Mus muscu | 560 | 16 | 1.6 343 | 59 | HV1S31784 | |
| 488 | 17 | 1.7 219293 | 72 | AC053546 | Homo sapi | 561 | 16 | 1.6 343 | 59 | HV1S31785 | |
| 489 | 17 | 1.7 220995 | 61 | AC010687 | Homo sapi | 562 | 16 | 1.6 343 | 59 | HV1S31786 | |
| 490 | 17 | 1.7 221913 | 67 | AC022129 | Homo sapi | 563 | 16 | 1.6 343 | 59 | HV1S31787 | |
| 491 | 17 | 1.7 223000 | 6 | LMF1CHR4A | Leishmani | 564 | 16 | 1.6 343 | 59 | HV1S3781 | |
| 492 | 17 | 1.7 224583 | 75 | AC074153 | Mus muscu | 565 | 16 | 1.6 343 | 59 | HV1S3781 | |
| 493 | 17 | 1.7 226060 | 76 | AC079583 | Mus muscu | 566 | 16 | 1.6 343 | 59 | HV1S3781 | |
| 494 | 17 | 1.7 226483 | 5 | AE003693 | Drosophill | 567 | 16 | 1.6 343 | 59 | HV1S3781 | |
| 495 | 17 | 1.7 227848 | 5 | AE003692 | Drosophill | 568 | 16 | 1.6 343 | 59 | HV1S3782 | |
| 496 | 17 | 1.7 228859 | 66 | AC021141 | Homo sapi | 569 | 16 | 1.6 343 | 59 | HV1S3783 | |
| 497 | 17 | 1.7 229122 | 5 | AE003700 | Drosophill | 570 | 16 | 1.6 343 | 59 | HV1S3784 | |
| 498 | 17 | 1.7 229252 | 81 | AL390732 | Homo sapi | 571 | 16 | 1.6 343 | 59 | HV1S3785 | |
| 499 | 17 | 1.7 229380 | 76 | AC079636 | Mus muscu | 572 | 16 | 1.6 343 | 59 | HV1S3786 | |
| 500 | 17 | 1.7 230651 | 74 | AC073535 | Homo sapi | 573 | 16 | 1.6 343 | 59 | HV1S3787 | |
| 501 | 17 | 1.7 233752 | 74 | AC073683 | Mus muscu | 574 | 16 | 1.6 343 | 59 | HV1S3788 | |
| 502 | 17 | 1.7 240285 | 60 | AC008939 | Homo sapi | 575 | 16 | 1.6 343 | 59 | HV1S3789 | |
| 503 | 17 | 1.7 240825 | 10 | AX087869 | Sequence | 576 | 16 | 1.6 343 | 59 | HV1PDU561 | |
| 504 | 17 | 1.7 241753 | 65 | AC019149 | Homo sapi | 577 | 16 | 1.6 343 | 59 | HV1PDU562 | |
| 505 | 17 | 1.7 242662 | 76 | AC080020 | Mus muscu | 578 | 16 | 1.6 344 | 58 | AF028581 | |
| 506 | 17 | 1.7 245880 | 97 | AC079387 | Homo sapi | 579 | 16 | 1.6 354 | 58 | AF0224030 | |
| 507 | 17 | 1.7 250529 | 96 | HUAE000658 | Homo sapi | 580 | 16 | 1.6 354 | 58 | AF020207 | Rattus no |
| 508 | 17 | 1.7 255004 | 80 | AL358232 | Homo sapi | 581 | 16 | 1.6 356 | 58 | AF050932 | |
| 509 | 17 | 1.7 258829 | 75 | AC073771 | Mus muscu | 582 | 16 | 1.6 357 | 58 | AF191457 | |
| 510 | 17 | 1.7 259421 | 74 | AC073689 | Mus muscu | 583 | 16 | 1.6 360 | 58 | AF073390 | HIV-1 iso |
| 511 | 17 | 1.7 260233 | 72 | AC055739 | Homo sapi | 584 | 16 | 1.6 360 | 58 | AF224046 | HIV-1 pat |
| 512 | 17 | 1.7 263190 | 63 | AC015657 | Mus muscu | 585 | 16 | 1.6 368 | 58 | AF017866 | HIV-1 iso |
| 513 | 17 | 1.7 283099 | 4 | AE003466 | Drosophill | 586 | 16 | 1.6 372 | 94 | AF050208 | Rattus no |
| 514 | 17 | 1.7 285479 | 72 | AC051621 | Mus muscu | 587 | 16 | 1.6 372 | 94 | AF050208 | Rattus no |
| 515 | 17 | 1.7 286758 | 86 | AC006449 | Homo sapi | 588 | 16 | 1.6 374 | 54 | G38490 | |
| 516 | 17 | 1.7 290581 | 75 | AC078960 | Homo sapi | 589 | 16 | 1.6 375 | 58 | AF050937 | HIV-1 iso |
| 517 | 17 | 1.7 291136 | 2 | AP002556 | Escherich | 590 | 16 | 1.6 380 | 58 | AF050939 | HIV-1 iso |
| 518 | 17 | 1.7 294308 | 4 | AE003595 | Drosophill | 591 | 16 | 1.6 380 | 58 | AF050935 | HIV-1 iso |
| 519 | 17 | 1.7 303250 | 2 | AP001518 | Bacillus | 592 | 16 | 1.6 385 | 58 | AF073402 | HIV-1 iso |

| | | | | | | | | | | | | | | | | |
|-----|----|-----|-----|----|----------|-----------|-------|-----|----|-----|-----|----|------------|------------|--------------|-------|
| 593 | 16 | 1.6 | 386 | 58 | AF028584 | HIV-1 | iso | 666 | 16 | 1.6 | 400 | 58 | AF174967 | AF174967 | HIV-1 | iso |
| 594 | 16 | 1.6 | 386 | 58 | AF028587 | HIV-1 | iso | 667 | 16 | 1.6 | 400 | 58 | AF174968 | AF174968 | HIV-1 | iso |
| 595 | 16 | 1.6 | 388 | 58 | AF014238 | HIV-1 | Pat | 668 | 16 | 1.6 | 400 | 58 | AF174969 | AF174969 | HIV-1 | iso |
| 596 | 16 | 1.6 | 388 | 58 | AF014239 | HIV-1 | Pat | 669 | 16 | 1.6 | 400 | 58 | AF174970 | AF174970 | HIV-1 | iso |
| 597 | 16 | 1.6 | 388 | 58 | AF014240 | HIV-1 | Pat | 670 | 16 | 1.6 | 400 | 58 | AF174971 | AF174971 | HIV-1 | iso |
| 598 | 16 | 1.6 | 388 | 58 | AF014344 | HIV-1 | Pat | 671 | 16 | 1.6 | 401 | 58 | AF028583 | AF028583 | HIV-1 | iso |
| 599 | 16 | 1.6 | 388 | 58 | AF014345 | HIV-1 | Pat | 672 | 16 | 1.6 | 401 | 58 | AF073401 | AF073401 | HIV-1 | iso |
| 600 | 16 | 1.6 | 388 | 58 | AF014358 | HIV-1 | Pat | 673 | 16 | 1.6 | 406 | 58 | AF073400 | AF073400 | HIV-1 | iso |
| 601 | 16 | 1.6 | 390 | 58 | AF015406 | HIV-1 | str | 674 | 16 | 1.6 | 409 | 58 | AF014214 | AF014214 | HIV-1 | Pat |
| 602 | 16 | 1.6 | 390 | 58 | AF062050 | HIV-1 | iso | 675 | 16 | 1.6 | 411 | 94 | AF020210 | AF020210 | Rattus no | |
| 603 | 16 | 1.6 | 391 | 58 | AF014194 | HIV-1 | Pat | 676 | 16 | 1.6 | 412 | 58 | AF014335 | AF014335 | HIV-1 | Pat |
| 604 | 16 | 1.6 | 391 | 58 | AF014199 | HIV-1 | Pat | 677 | 16 | 1.6 | 414 | 4 | AB040074 | AB040074 | Hydra mag | |
| 605 | 16 | 1.6 | 391 | 58 | AF014202 | HIV-1 | Pat | 678 | 16 | 1.6 | 414 | 58 | AF028577 | AF028577 | HIV-1 | iso |
| 606 | 16 | 1.6 | 391 | 58 | AF014216 | HIV-1 | Pat | 679 | 16 | 1.6 | 416 | 97 | HUM0293F03 | HUM0293F03 | | |
| 607 | 16 | 1.6 | 391 | 58 | AF014225 | HIV-1 | Pat | 680 | 16 | 1.6 | 420 | 58 | AF028580 | AF028580 | HIV-1 | iso |
| 608 | 16 | 1.6 | 391 | 58 | AF014228 | HIV-1 | Pat | 681 | 16 | 1.6 | 422 | 54 | G65449 | G65449 | SMCY-M81 | Ra |
| 609 | 16 | 1.6 | 391 | 58 | AF014233 | HIV-1 | Pat | 682 | 16 | 1.6 | 448 | 58 | AF073403 | AF073403 | HIV-1 | iso |
| 610 | 16 | 1.6 | 391 | 58 | AF050899 | HIV-1 | mat | 683 | 16 | 1.6 | 452 | 8 | AF076295 | AF076295 | Lampbrush | |
| 611 | 16 | 1.6 | 391 | 58 | AF050900 | HIV-1 | mat | 684 | 16 | 1.6 | 454 | 58 | AF028586 | AF028586 | HIV-1 | iso |
| 612 | 16 | 1.6 | 391 | 58 | AF050927 | HIV-1 | iso | 685 | 16 | 1.6 | 480 | 14 | CNS01BXR | CNS01BXR | | |
| 613 | 16 | 1.6 | 391 | 58 | AF050928 | HIV-1 | iso | 686 | 16 | 1.6 | 492 | 3 | SCSCYEEN | SCSCYEEN | | |
| 614 | 16 | 1.6 | 391 | 58 | AF050929 | HIV-1 | iso | 687 | 16 | 1.6 | 492 | 3 | SGSGHE | SGSGHE | | |
| 615 | 16 | 1.6 | 391 | 58 | AF050931 | HIV-1 | iso | 688 | 16 | 1.6 | 492 | 3 | SGSGLEGEN | SGSGLEGEN | | |
| 616 | 16 | 1.6 | 391 | 58 | AF050936 | HIV-1 | iso | 689 | 16 | 1.6 | 504 | 97 | S63672 | S63672 | trichohyal | |
| 617 | 16 | 1.6 | 391 | 58 | AF050942 | HIV-1 | iso | 690 | 16 | 1.6 | 523 | 97 | HUM0285H04 | HUM0285H04 | | |
| 618 | 16 | 1.6 | 391 | 58 | AF050944 | HIV-1 | iso | 691 | 16 | 1.6 | 541 | 9 | RP038304 | RP038304 | Roscoe | P |
| 619 | 16 | 1.6 | 391 | 58 | AF050946 | HIV-1 | iso | 692 | 16 | 1.6 | 559 | 14 | CSP388301 | CSP388301 | Caulleya | |
| 620 | 16 | 1.6 | 391 | 58 | AF050951 | HIV-1 | iso | 693 | 16 | 1.6 | 561 | 14 | HM0388297 | HM0388297 | Hedychium | |
| 621 | 16 | 1.6 | 391 | 58 | AF050952 | HIV-1 | iso | 694 | 16 | 1.6 | 562 | 14 | CGR388299 | CGR388299 | Caulleya | |
| 622 | 16 | 1.6 | 391 | 58 | AF050953 | HIV-1 | iso | 695 | 16 | 1.6 | 563 | 14 | HFU388296 | HFU388296 | Hedychium | |
| 623 | 16 | 1.6 | 391 | 58 | AF050959 | HIV-1 | iso | 696 | 16 | 1.6 | 565 | 14 | HS421120S | HS421120S | H. sapientis | |
| 624 | 16 | 1.6 | 391 | 58 | AF050960 | HIV-1 | iso | 697 | 16 | 1.6 | 574 | 54 | AX034339 | AX034339 | Sequence | |
| 625 | 16 | 1.6 | 391 | 58 | AF051008 | HIV-1 | iso | 698 | 16 | 1.6 | 598 | 5 | AF320100 | AF320100 | Dendronep | |
| 626 | 16 | 1.6 | 391 | 58 | AF051009 | HIV-1 | iso | 699 | 16 | 1.6 | 602 | 5 | AF320102 | AF320102 | Dendronep | |
| 627 | 16 | 1.6 | 391 | 58 | AF051010 | HIV-1 | iso | 700 | 16 | 1.6 | 603 | 5 | AF320103 | AF320103 | Dendronep | |
| 628 | 16 | 1.6 | 391 | 58 | AF051012 | HIV-1 | iso | 701 | 16 | 1.6 | 603 | 5 | AF320109 | AF320109 | Dendronep | |
| 629 | 16 | 1.6 | 391 | 58 | AF051013 | HIV-1 | iso | 702 | 16 | 1.6 | 603 | 5 | AF3202405 | AF3202405 | Pommereesc | |
| 630 | 16 | 1.6 | 391 | 58 | AF051014 | HIV-1 | iso | 703 | 16 | 1.6 | 604 | 5 | AF3202380 | AF3202380 | Hedychium | |
| 631 | 16 | 1.6 | 391 | 58 | AF051016 | HIV-1 | iso | 704 | 16 | 1.6 | 604 | 5 | AF320101 | AF320101 | Dendronep | |
| 632 | 16 | 1.6 | 391 | 58 | AF051019 | HIV-1 | iso | 705 | 16 | 1.6 | 608 | 54 | G39591 | G39591 | Zebra | |
| 633 | 16 | 1.6 | 391 | 58 | AF051022 | HIV-1 | iso | 706 | 16 | 1.6 | 610 | 5 | AF320099 | AF320099 | Dendronep | |
| 634 | 16 | 1.6 | 391 | 58 | AF051024 | HIV-1 | iso | 707 | 16 | 1.6 | 612 | 13 | AF202374 | AF202374 | Hedychium | |
| 635 | 16 | 1.6 | 391 | 58 | AF051027 | HIV-1 | iso | 708 | 16 | 1.6 | 621 | 13 | AF202403 | AF202403 | Hedychium | |
| 636 | 16 | 1.6 | 391 | 58 | AF051030 | HIV-1 | iso | 709 | 16 | 1.6 | 621 | 13 | AF202405 | AF202405 | Pommereesc | |
| 637 | 16 | 1.6 | 393 | 94 | AF020209 | Rattus no | | 710 | 16 | 1.6 | 623 | 13 | AF202419 | AF202419 | Caulleya | |
| 638 | 16 | 1.6 | 394 | 58 | AF014269 | HIV-1 | Pat | 711 | 16 | 1.6 | 626 | 13 | AF202385 | AF202385 | Hedychium | |
| 639 | 16 | 1.6 | 396 | 58 | AF084766 | HIV-1 | Pat | 712 | 16 | 1.6 | 626 | 13 | AF202388 | AF202388 | Hedychium | |
| 640 | 16 | 1.6 | 396 | 58 | AF084767 | HIV-1 | Pat | 713 | 16 | 1.6 | 626 | 13 | AF202395 | AF202395 | Hedychium | |
| 641 | 16 | 1.6 | 396 | 58 | AF084768 | HIV-1 | Pat | 714 | 16 | 1.6 | 626 | 13 | AF202406 | AF202406 | Roscoe | h |
| 642 | 16 | 1.6 | 396 | 58 | AF084769 | HIV-1 | Pat | 715 | 16 | 1.6 | 627 | 13 | AF202376 | AF202376 | Hedychium | |
| 643 | 16 | 1.6 | 396 | 58 | AF084770 | HIV-1 | Pat | 716 | 16 | 1.6 | 627 | 13 | AF202380 | AF202380 | Hedychium | |
| 644 | 16 | 1.6 | 396 | 58 | AF084771 | HIV-1 | Pat | 717 | 16 | 1.6 | 627 | 13 | AF202381 | AF202381 | Hedychium | |
| 645 | 16 | 1.6 | 396 | 58 | AF084772 | HIV-1 | Pat | 718 | 16 | 1.6 | 627 | 13 | AF202382 | AF202382 | Hedychium | |
| 646 | 16 | 1.6 | 396 | 58 | AF084773 | HIV-1 | Pat | 719 | 16 | 1.6 | 627 | 13 | AF202393 | AF202393 | Hedychium | |
| 647 | 16 | 1.6 | 396 | 58 | AF084774 | HIV-1 | Pat | 720 | 16 | 1.6 | 627 | 13 | AF202394 | AF202394 | Hedychium | |
| 648 | 16 | 1.6 | 396 | 58 | AF084775 | HIV-1 | Pat | 721 | 16 | 1.6 | 627 | 13 | AF202397 | AF202397 | Hedychium | |
| 649 | 16 | 1.6 | 396 | 58 | AF084776 | HIV-1 | Pat | 722 | 16 | 1.6 | 628 | 13 | AF202377 | AF202377 | Hedychium | |
| 650 | 16 | 1.6 | 396 | 58 | AF084777 | HIV-1 | Pat | 723 | 16 | 1.6 | 628 | 13 | AF202378 | AF202378 | Hedychium | |
| 651 | 16 | 1.6 | 396 | 59 | HIV1DUMC | Human | immun | 724 | 16 | 1.6 | 628 | 13 | AF202383 | AF202383 | Hedychium | |
| 652 | 16 | 1.6 | 396 | 59 | HIV1DUMC | Human | immun | 725 | 16 | 1.6 | 628 | 13 | AF202384 | AF202384 | Hedychium | |
| 653 | 16 | 1.6 | 397 | 58 | AF014195 | HIV-1 | Pat | 726 | 16 | 1.6 | 628 | 13 | AF202389 | AF202389 | Hedychium | |
| 654 | 16 | 1.6 | 399 | 58 | AF174897 | HIV-1 | iso | 727 | 16 | 1.6 | 628 | 13 | AF202390 | AF202390 | Hedychium | |
| 655 | 16 | 1.6 | 399 | 58 | AF174898 | HIV-1 | iso | 728 | 16 | 1.6 | 628 | 13 | AF202391 | AF202391 | Hedychium | |
| 656 | 16 | 1.6 | 399 | 58 | AF174899 | HIV-1 | iso | 729 | 16 | 1.6 | 628 | 13 | AF202392 | AF202392 | Hedychium | |
| 657 | 16 | 1.6 | 400 | 58 | AF014222 | HIV-1 | Pat | 730 | 16 | 1.6 | 628 | 13 | AF202396 | AF202396 | Hedychium | |
| 658 | 16 | 1.6 | 400 | 58 | AF174851 | HIV-1 | iso | 731 | 16 | 1.6 | 629 | 13 | AF202379 | AF202379 | Hedychium | |
| 659 | 16 | 1.6 | 400 | 58 | AF174855 | HIV-1 | iso | 732 | 16 | 1.6 | 629 | 13 | AF202402 | AF202402 | Hedychium | |
| 660 | 16 | 1.6 | 400 | 58 | AF174896 | HIV-1 | iso | 733 | 16 | 1.6 | 631 | 13 | AF202401 | AF202401 | Hedychium | |
| 661 | 16 | 1.6 | 400 | 58 | AF174902 | HIV-1 | iso | 734 | 16 | 1.6 | 631 | 13 | AF202401 | AF202401 | Hedychium | |
| 662 | 16 | 1.6 | 400 | 58 | AF174962 | HIV-1 | iso | 735 | 16 | 1.6 | 659 | 97 | HUM02858F | HUM02858F | Human | Cpg 1 |
| 663 | 16 | 1.6 | 400 | 58 | AF174964 | HIV-1 | iso | 736 | 16 | 1.6 | 671 | 8 | AF221933 | AF221933 | Musculus | |
| 664 | 16 | 1.6 | 400 | 58 | AF174965 | HIV-1 | iso | 737 | 16 | 1.6 | 704 | 54 | AB056582 | AB056582 | Xenopus 1 | |
| 665 | 16 | 1.6 | 400 | 58 | AF174966 | HIV-1 | iso | 738 | 16 | 1.6 | 720 | 9 | G61829 | G61829 | SHGC-89440 | |
| | | | | | | | | | | | | | AX065817 | AX065817 | Sequence | |

| | | | | | | | | | | | | | |
|-----|----|-----|------|----|-----------|----------------------------|-----|----|-----|------|----|-----------|------------------------|
| 739 | 16 | 1.6 | 729 | 58 | AF071291 | AF071291 HIV-1 iso | 812 | 16 | 1.6 | 1669 | 94 | AF273686 | AF273686 Mus muscu |
| 740 | 16 | 1.6 | 729 | 58 | AF071300 | AF071300 HIV-1 iso | 813 | 16 | 1.6 | 1669 | 94 | AF273687 | AF273687 Mus muscu |
| 741 | 16 | 1.6 | 729 | 58 | AF196682 | AF196682 HIV-1 iso | 814 | 16 | 1.6 | 1671 | 2 | AX024351 | AX024351 Sequence |
| 742 | 16 | 1.6 | 737 | 9 | AB6380 | AB6380 Sequence 10 | 815 | 16 | 1.6 | 1671 | 9 | AX024244 | AX024244 Sequence |
| 743 | 16 | 1.6 | 737 | 10 | E66398 | E66398 Genome DNA | 816 | 16 | 1.6 | 1699 | 3 | SABLA | E09684 |
| 744 | 16 | 1.6 | 738 | 58 | AF071302 | AF071302 HIV-1 iso | 817 | 16 | 1.6 | 1722 | 45 | AF273688 | AF273688 cDNA encodi |
| 745 | 16 | 1.6 | 738 | 58 | AF071308 | AF071308 HIV-1 iso | 818 | 16 | 1.6 | 1723 | 94 | AF273688 | AF273688 Mus muscu |
| 746 | 16 | 1.6 | 759 | 96 | TV4250739 | TV4250739 Trypanoso | 819 | 16 | 1.6 | 1734 | 45 | E11874 | E11874 Genomic DNA |
| 747 | 16 | 1.6 | 774 | 14 | MA293103 | MA293103 Sequence 25 | 820 | 16 | 1.6 | 1780 | 89 | AK023355 | AK023355 Homo sapi |
| 748 | 16 | 1.6 | 774 | 14 | MA293103 | MA293103 Sequence 25 | 821 | 16 | 1.6 | 1788 | 93 | HSHP512 | HSHP512 Homo sapi |
| 749 | 16 | 1.6 | 785 | 6 | DMDBX1 | DMDBX1 D. melanog | 822 | 16 | 1.6 | 1790 | 12 | AF118126 | AF118126 Magnapor |
| 750 | 16 | 1.6 | 785 | 6 | DMDBX1 | DMDBX1 D. melanog | 823 | 16 | 1.6 | 1803 | 10 | AX090436 | AX090436 Sequence |
| 751 | 16 | 1.6 | 790 | 53 | CNS07270 | CNS07270 Clone BAO | 824 | 16 | 1.6 | 1810 | 9 | AX047400 | AX047400 Sequence |
| 752 | 16 | 1.6 | 849 | 53 | CNS07270 | CNS07270 Clone BAO | 825 | 16 | 1.6 | 1834 | 93 | HSMB00543 | HSMB00543 Sequence |
| 753 | 16 | 1.6 | 863 | 91 | BC005052 | BC005052 Homo sapi | 826 | 16 | 1.6 | 1839 | 12 | AK045892 | AK045892 Sequence |
| 754 | 16 | 1.6 | 863 | 91 | BC005052 | BC005052 Homo sapi | 827 | 16 | 1.6 | 1845 | 89 | AF308303 | AF308303 Homo sapi |
| 755 | 16 | 1.6 | 886 | 53 | CNS0750E | CNS0750E Homo sapi | 828 | 16 | 1.6 | 1864 | 9 | AR076171 | AR076171 Sequence |
| 756 | 16 | 1.6 | 892 | 72 | AC052426 | AC052426 Homo sapi | 829 | 16 | 1.6 | 1864 | 9 | AR076171 | AR076171 Sequence |
| 757 | 16 | 1.6 | 900 | 88 | AF077206 | AF077206 Homo sapi | 830 | 16 | 1.6 | 1864 | 12 | AF036949 | AF036949 Zea mays |
| 758 | 16 | 1.6 | 901 | 91 | BC005156 | BC005156 Homo sapi | 831 | 16 | 1.6 | 1870 | 14 | CIN242538 | CIN242538 Sequence |
| 759 | 16 | 1.6 | 902 | 93 | HSMB00572 | HSMB00572 Homo sapi | 832 | 16 | 1.6 | 1900 | 94 | AF287300 | AF287300 Rattus no |
| 760 | 16 | 1.6 | 904 | 91 | BC005393 | BC005393 Homo sapi | 833 | 16 | 1.6 | 1907 | 10 | AX086887 | AX086887 Sequence |
| 761 | 16 | 1.6 | 910 | 75 | AC075071 | AC075071 Giardia i | 834 | 16 | 1.6 | 1984 | 4 | AB012617 | AB012617 Caenorhab |
| 762 | 16 | 1.6 | 912 | 53 | CNS06KRW | CNS06KRW T7 end of | 835 | 16 | 1.6 | 2025 | 3 | RE041754 | RE041754 Rhizobium e |
| 763 | 16 | 1.6 | 926 | 53 | CNS07308 | CNS07308 T7 end of | 836 | 16 | 1.6 | 2027 | 94 | AB041037 | AB041037 Mus muscu |
| 764 | 16 | 1.6 | 926 | 53 | CNS07308 | CNS07308 T7 end of | 837 | 16 | 1.6 | 2037 | 2 | CGGDH | CGGDH |
| 765 | 16 | 1.6 | 948 | 53 | CNS06G78 | CNS06G78 T7 end of | 838 | 16 | 1.6 | 2037 | 2 | CGGDH | CGGDH |
| 766 | 16 | 1.6 | 952 | 76 | CNS06G78 | CNS06G78 T7 end of | 839 | 16 | 1.6 | 2040 | 9 | AX009994 | AX009994 Sequence |
| 767 | 16 | 1.6 | 952 | 76 | CNS06G78 | CNS06G78 T7 end of | 840 | 16 | 1.6 | 2045 | 94 | AB016229 | AB016229 Mus muscu |
| 768 | 16 | 1.6 | 966 | 53 | CNS07630 | CNS07630 Clone XBA | 841 | 16 | 1.6 | 2047 | 94 | AB016230 | AB016230 Mus muscu |
| 769 | 16 | 1.6 | 981 | 8 | AF317286 | AF317286 Gallus ga | 842 | 16 | 1.6 | 2063 | 9 | AX048734 | AX048734 Sequence |
| 770 | 16 | 1.6 | 990 | 10 | AX089467 | AX089467 Sequence | 843 | 16 | 1.6 | 2065 | 88 | AF049459 | AF049459 Homo sapi |
| 771 | 16 | 1.6 | 993 | 93 | HSMB01069 | HSMB01069 Giardia i | 844 | 16 | 1.6 | 2076 | 88 | AF075682 | AF075682 Hemorrhag |
| 772 | 16 | 1.6 | 999 | 74 | AC070320 | AC070320 Giardia i | 845 | 16 | 1.6 | 2089 | 89 | AF026182 | AF026182 Homo sapi |
| 773 | 16 | 1.6 | 1016 | 53 | CNS06KBA | CNS06KBA T3 end of | 846 | 16 | 1.6 | 2140 | 97 | H0MTNEBA | H0MTNEBA Human tumor |
| 774 | 16 | 1.6 | 1022 | 76 | AC080628 | AC080628 Giardia i | 847 | 16 | 1.6 | 2159 | 89 | AK000953 | AK000953 Homo sapi |
| 775 | 16 | 1.6 | 1024 | 72 | AC060129 | AC060129 Trachemys | 848 | 16 | 1.6 | 2163 | 94 | BC004039 | BC004039 Mus muscu |
| 776 | 16 | 1.6 | 1047 | 53 | CNS06LYS | AC060129 Giardia i | 849 | 16 | 1.6 | 2184 | 94 | AF132727 | AF132727 Rattus no |
| 777 | 16 | 1.6 | 1047 | 53 | CNS06LYS | AC060129 Giardia i | 850 | 16 | 1.6 | 2192 | 15 | SCVNL253W | SCVNL253W S. cerevisia |
| 778 | 16 | 1.6 | 1061 | 53 | CNS0738M | AL427196 clone BAO | 851 | 16 | 1.6 | 2199 | 95 | RATSE | RATSE Rat mRNA fo |
| 779 | 16 | 1.6 | 1081 | 94 | AB018562 | AB018562 Mus muscu | 852 | 16 | 1.6 | 2214 | 9 | AR093957 | AR093957 Sequence |
| 780 | 16 | 1.6 | 1150 | 94 | OCU77896 | OCU77896 Oryctolagus | 853 | 16 | 1.6 | 2214 | 10 | E15639 | E15639 Pinctada fu |
| 781 | 16 | 1.6 | 1159 | 94 | AF148849 | AF148849 Mus muscu | 854 | 16 | 1.6 | 2226 | 2 | AX024347 | AX024347 Sequence |
| 782 | 16 | 1.6 | 1208 | 91 | BC001852 | BC001852 Homo sapi | 855 | 16 | 1.6 | 2226 | 9 | AX024240 | AX024240 Sequence |
| 783 | 16 | 1.6 | 1216 | 93 | HSMB01637 | HSMB01637 Homo sapi | 856 | 16 | 1.6 | 2232 | 88 | AF049460 | AF049460 Sequence |
| 784 | 16 | 1.6 | 1233 | 89 | AK000553 | AK000553 Homo sapi | 857 | 16 | 1.6 | 2239 | 2 | AF270371 | AF270371 Staphyloc |
| 785 | 16 | 1.6 | 1247 | 85 | AB051152 | AB051152 Macaca fa | 858 | 16 | 1.6 | 2247 | 94 | MHT000001 | MHT000001 Mus muscu |
| 786 | 16 | 1.6 | 1250 | 94 | MMCD40 | MMCD40 M. musculus | 859 | 16 | 1.6 | 2255 | 6 | DMREFPNN | DMREFPNN Mus muscu |
| 787 | 16 | 1.6 | 1260 | 94 | AF202778 | AF202778 Mus muscu | 860 | 16 | 1.6 | 2263 | 89 | AK021663 | AK021663 Staphyloc |
| 788 | 16 | 1.6 | 1292 | 5 | UBU85123 | UBU85123 Unidentifie | 861 | 16 | 1.6 | 2271 | 96 | TRBEPCK | TRBEPCK Trypanosoma |
| 789 | 16 | 1.6 | 1299 | 5 | AF017268 | AF017268 Toxoplasma | 862 | 16 | 1.6 | 2274 | 2 | AX024349 | AX024349 Sequence |
| 790 | 16 | 1.6 | 1314 | 9 | AR072866 | AR072866 Sequence | 863 | 16 | 1.6 | 2274 | 9 | AX024242 | AX024242 Sequence |
| 791 | 16 | 1.6 | 1314 | 10 | E14942 | E14942 Aspergillus | 864 | 16 | 1.6 | 2292 | 94 | AF006071 | AF006071 Mus muscu |
| 792 | 16 | 1.6 | 1314 | 10 | E14955 | E14955 Aspergillus | 865 | 16 | 1.6 | 2324 | 89 | AF148713 | AF148713 Homo sapi |
| 793 | 16 | 1.6 | 1314 | 10 | E14955 | E14955 Aspergillus | 866 | 16 | 1.6 | 2324 | 88 | AF049460 | AF049460 Sequence |
| 794 | 16 | 1.6 | 1314 | 13 | ATFAOA | ATFAOA Y09020 A. terreus m | 867 | 16 | 1.6 | 2332 | 3 | STMHRDB | STMHRDB Streptomyces |
| 795 | 16 | 1.6 | 1381 | 94 | AF273690 | AF273690 Mus muscu | 868 | 16 | 1.6 | 2335 | 94 | DA33961 | DA33961 Mus muscu |
| 796 | 16 | 1.6 | 1381 | 94 | MMU77630 | MMU77630 Mus muscu | 869 | 16 | 1.6 | 2335 | 94 | DA33961 | DA33961 Mus muscu |
| 797 | 16 | 1.6 | 1417 | 96 | TEU64517 | TEU64517 Toxoplasma | 870 | 16 | 1.6 | 2344 | 7 | RABPKC | RABPKC CRAUJ765 |
| 798 | 16 | 1.6 | 1449 | 89 | AF287302 | AF287302 Homo sapi | 871 | 16 | 1.6 | 2361 | 10 | GBGEN | GBGEN Xenopus lae |
| 799 | 16 | 1.6 | 1455 | 97 | IR1845189 | IR1845189 Homo sapi | 872 | 16 | 1.6 | 2363 | 10 | DMREFPCK | DMREFPCK Sequence |
| 800 | 16 | 1.6 | 1464 | 9 | AX063811 | AX063811 Sequence | 873 | 16 | 1.6 | 2394 | 88 | AF061795 | AF061795 Homo sapi |
| 801 | 16 | 1.6 | 1468 | 3 | TTASNS | TTASNS X91009 T. thermophi | 874 | 16 | 1.6 | 2416 | 91 | BC000770 | BC000770 Homo sapi |
| 802 | 16 | 1.6 | 1468 | 89 | AF119879 | AF119879 Homo sapi | 875 | 16 | 1.6 | 2416 | 91 | BC000770 | BC000770 Homo sapi |
| 803 | 16 | 1.6 | 1477 | 59 | HIV1205 | HIV1205 Human Immun | 876 | 16 | 1.6 | 2449 | 85 | AB006965 | AB006965 Homo sapi |
| 804 | 16 | 1.6 | 1513 | 10 | AX082314 | AX082314 Sequence | 877 | 16 | 1.6 | 2452 | 94 | AF273685 | AF273685 Mus muscu |
| 805 | 16 | 1.6 | 1519 | 94 | AF273689 | AF273689 Mus muscu | 878 | 16 | 1.6 | 2472 | 89 | AF151685 | AF151685 Homo sapi |
| 806 | 16 | 1.6 | 1525 | 94 | AB005049 | AB005049 Cavia por | 879 | 16 | 1.6 | 2480 | 94 | MUSEMTX | MUSEMTX Mus muscu |
| 807 | 16 | 1.6 | 1543 | 94 | AF156814 | AF156814 Cavia por | 880 | 16 | 1.6 | 2487 | 94 | AF075533 | AF075533 Mus muscu |
| 808 | 16 | 1.6 | 1581 | 94 | AF134488 | AF134488 Cavia por | 881 | 16 | 1.6 | 2493 | 88 | AF000430 | AF000430 Homo sapi |
| 809 | 16 | 1.6 | 1600 | 3 | RCAN1FO | M26523 Rhododactyl | 882 | 16 | 1.6 | 2512 | 8 | XLU68724 | XLU68724 Xenopus lae |
| 810 | 16 | 1.6 | 1662 | 94 | AB016423 | AB016423 Mus muscu | 883 | 16 | 1.6 | 2529 | 1 | AF028831 | AF028831 Cenarchae |
| 811 | 16 | 1.6 | 1668 | 8 | AF258786 | AF258786 Danio rer | 884 | 16 | 1.6 | 2544 | 58 | AF037263 | AF037263 Spodopter |

| | | | | | | |
|-------|----|-----|------|----|-----------|---------------------|
| C 885 | 16 | 1.6 | 2545 | 59 | SLI011858 | AJ011858 Spodopter |
| C 886 | 16 | 1.6 | 2610 | 9 | AX023369 | AJ023369 Sequence |
| C 887 | 16 | 1.6 | 2610 | 95 | RNAJ3148 | AJ003148 Rattus no |
| C 888 | 16 | 1.6 | 2632 | 94 | AF273684 | AF273684 Mus muscu |
| C 889 | 16 | 1.6 | 2661 | 15 | TRAGLUCS | 268706 T. reeset mr |
| C 890 | 16 | 1.6 | 2673 | 14 | PAM133148 | AJ133148 Persea am |
| C 891 | 16 | 1.6 | 2684 | 6 | DMU23926 | AJ3926 Drosophila |
| C 892 | 16 | 1.6 | 2689 | 1 | AF128272 | AF128272 Streptomy |
| C 893 | 16 | 1.6 | 2691 | 6 | DMU23923 | DMU23923 Drosophila |
| C 894 | 16 | 1.6 | 2691 | 6 | DMU23927 | DMU23927 Drosophila |
| C 895 | 16 | 1.6 | 2691 | 6 | DMU23930 | DMU23930 Drosophila |
| C 896 | 16 | 1.6 | 2707 | 6 | DMU23930 | DMU23930 Drosophila |
| C 897 | 16 | 1.6 | 2708 | 6 | DMU23928 | DMU23928 Drosophila |
| C 898 | 16 | 1.6 | 2708 | 6 | DMU23932 | DMU23932 Drosophila |
| C 899 | 16 | 1.6 | 2709 | 6 | DMU23929 | DMU23929 Drosophila |
| C 900 | 16 | 1.6 | 2719 | 6 | DMU23924 | DMU23924 Drosophila |
| C 901 | 16 | 1.6 | 2719 | 6 | DMU23925 | DMU23925 Drosophila |
| C 902 | 16 | 1.6 | 2720 | 6 | AK002127 | AK002127 Homo sapi |
| C 903 | 16 | 1.6 | 2772 | 89 | BC004237 | BC004237 Homo sapi |
| C 904 | 16 | 1.6 | 2776 | 91 | BC004237 | BC004237 Homo sapi |
| C 905 | 16 | 1.6 | 2795 | 95 | RNU78889 | U78889 Rattus norv |
| C 906 | 16 | 1.6 | 2823 | 2 | ECOHSD | L18758 Escherichia |
| C 907 | 16 | 1.6 | 2891 | 9 | AX065435 | AX065435 Sequence |
| C 908 | 16 | 1.6 | 2916 | 8 | CHKMPOLE | D28598 Chicken mRN |
| C 909 | 16 | 1.6 | 2932 | 93 | HSMB01716 | AL136748 Homo sapi |
| C 910 | 16 | 1.6 | 2955 | 58 | AF020660 | AF020660 Tomato sp |
| C 911 | 16 | 1.6 | 3045 | 9 | A96458 | A96458 Sequence 49 |
| C 912 | 16 | 1.6 | 3045 | 9 | A96462 | AF273683 Mus muscu |
| C 913 | 16 | 1.6 | 3055 | 94 | AF273683 | E15417 Arthropacte |
| C 914 | 16 | 1.6 | 3153 | 10 | E15417 | U51337 Streptomyce |
| C 915 | 16 | 1.6 | 3215 | 3 | SCU51332 | BC000136 Homo sapi |
| C 916 | 16 | 1.6 | 3226 | 91 | BC000136 | X66022 R. norvegicu |
| C 917 | 16 | 1.6 | 3233 | 95 | RNND4P | Z69085 Human DNA s |
| C 918 | 16 | 1.6 | 3258 | 92 | HS361H4B | AF006675 Rhodother |
| C 919 | 16 | 1.6 | 3290 | 6 | DB6074 | DB6074 Pinctada fu |
| C 920 | 16 | 1.6 | 3331 | 9 | AR039358 | AR039358 Sequence |
| C 921 | 16 | 1.6 | 3331 | 9 | AR039358 | AR039358 Sequence |
| C 922 | 16 | 1.6 | 3331 | 10 | E15640 | E15640 Pinctada fu |
| C 923 | 16 | 1.6 | 3351 | 91 | BC001211 | BC001211 Homo sapi |
| C 924 | 16 | 1.6 | 3381 | 8 | S63276 | S63276 SC1-100 kda |
| C 925 | 16 | 1.6 | 3444 | 8 | CHKMYLCC | J03397 Chicken myo |
| C 926 | 16 | 1.6 | 3453 | 3 | DMU76378 | X57478 Drosophila |
| C 927 | 16 | 1.6 | 3492 | 3 | RCSOR | X57478 Rhodobacter |
| C 928 | 16 | 1.6 | 3500 | 3 | SEORF1 | L37354 Saccharopol |
| C 929 | 16 | 1.6 | 3539 | 94 | D78349 | D78349 Mouse DNA f |
| C 930 | 16 | 1.6 | 3577 | 3 | RNDNAB | Y13813 Rhodothermu |
| C 931 | 16 | 1.6 | 3600 | 14 | NEURDHA | L13964 Neurospora |
| C 932 | 16 | 1.6 | 3621 | 85 | AB033116 | AB033116 Homo sapi |
| C 933 | 16 | 1.6 | 3682 | 89 | AK001713 | AK001713 Homo sapi |
| C 934 | 16 | 1.6 | 3747 | 14 | BEPA71 | AJ001045 Blastocla |
| C 935 | 16 | 1.6 | 3750 | 14 | CRU279021 | AJ279021 Candida r |
| C 936 | 16 | 1.6 | 3755 | 14 | AC008239 | AC008239 Leishmani |
| C 937 | 16 | 1.6 | 3845 | 6 | AF019043 | AF019043 Rattus no |
| C 938 | 16 | 1.6 | 3892 | 94 | DEREP2PE | X69831 D. erecta re |
| C 939 | 16 | 1.6 | 3915 | 3 | KPU00985 | U00985 Klebsiella |
| C 940 | 16 | 1.6 | 3952 | 3 | PSEDCD | J05282 P. cepacia 2 |
| C 941 | 16 | 1.6 | 3957 | 91 | AY009106 | AY009106 Homo sapi |
| C 942 | 16 | 1.6 | 3960 | 6 | CHKMLF | D28598 Chicken gen |
| C 943 | 16 | 1.6 | 3984 | 6 | DMREF2POL | X16993 Drosophila |
| C 944 | 16 | 1.6 | 4088 | 94 | MUSB3RP | J04036 Mus musculu |
| C 945 | 16 | 1.6 | 4107 | 59 | MOCREPEQB | M98815 Molluscum c |
| C 946 | 16 | 1.6 | 4231 | 94 | MUSYRKIN | D14042 Mus musculu |
| C 947 | 16 | 1.6 | 4261 | 1 | AB016241 | AB016241 Streptomy |
| C 948 | 16 | 1.6 | 4265 | 1 | AB028210 | AB028210 Streptomy |
| C 949 | 16 | 1.6 | 4294 | 94 | MUSITKA | L00619 Mouse T-cell |
| C 950 | 16 | 1.6 | 4394 | 2 | DB4102 | DB4102 Corynebacte |
| C 951 | 16 | 1.6 | 4394 | 2 | AR083912 | AR083912 Sequence |
| C 952 | 16 | 1.6 | 4412 | 10 | E14601 | E14601 Corynebacte |
| C 953 | 16 | 1.6 | 4494 | 13 | AF335465 | AF335465 Aspergill |
| C 954 | 16 | 1.6 | | | | |
| C 955 | 16 | 1.6 | | | | |
| C 956 | 16 | 1.6 | | | | |
| C 957 | 16 | 1.6 | | | | |

| | | | | | | |
|--------|----|-----|-------|----|-------------|----------------------|
| C 958 | 16 | 1.6 | 4525 | 3 | PFL1233 | Y12268 P. fluoresce |
| C 959 | 16 | 1.6 | 4543 | 6 | DMUBXGENE | X76220 D.melanogas |
| C 960 | 16 | 1.6 | 4589 | 7 | RABABALG | M93344 Rabbit acid |
| C 961 | 16 | 1.6 | 4648 | 63 | AC014057 | AC014057 Drosophi |
| C 962 | 16 | 1.6 | 4736 | 13 | AF205944 | AF205944 Blastocla |
| C 963 | 16 | 1.6 | 4746 | 1 | AF034831 | AF034831 Rhizobium |
| C 964 | 16 | 1.6 | 4756 | 97 | HUMPTPS | D37781 Human mRNA |
| C 965 | 16 | 1.6 | 4823 | 89 | AK024484 | AK024484 Homo sapi |
| C 966 | 16 | 1.6 | 4887 | 5 | AF531187 | AF531187 Drosophi |
| C 967 | 16 | 1.6 | 4972 | 3 | MAMAMIRM | X79027 M. mammonlaph |
| C 968 | 16 | 1.6 | 5084 | 3 | SGSTRELMB | X62567 S. griseus s |
| C 969 | 16 | 1.6 | 5117 | 9 | ARI09903 | ARI09903 Sequence |
| C 970 | 16 | 1.6 | 5117 | 97 | HSU10886 | U10886 Homo densi |
| C 971 | 16 | 1.6 | 5282 | 97 | HSY13034 | Y13034 Homo sapien |
| C 972 | 16 | 1.6 | 5407 | 2 | AF324349 | AF324349 Yersinia |
| C 973 | 16 | 1.6 | 5591 | 94 | MSMKELINA | Z22866 M. musculus |
| C 974 | 16 | 1.6 | 5596 | 1 | AF026544 | AF026544 Ralstonia |
| C 975 | 16 | 1.6 | 5645 | 94 | MMU012072 | AJ012072 Mus muscu |
| C 976 | 16 | 1.6 | 5774 | 94 | AF064748 | AF064748 Mus muscu |
| C 977 | 16 | 1.6 | 5994 | 3 | SVSNDDE | Y11547 S. virginiae |
| C 978 | 16 | 1.6 | 6213 | 2 | ASAJ187 | AJ000187 Arthropac |
| C 979 | 16 | 1.6 | 6396 | 7 | BTU02285 | U02285 Bos taurus |
| C 980 | 16 | 1.6 | 6412 | 95 | RNO131902 | AJ131902 Rattus no |
| C 981 | 16 | 1.6 | 6545 | 2 | ECOMALB | J01648 E. coli malB |
| C 982 | 16 | 1.6 | 6648 | 94 | AB015790 | AB015790 Mus muscu |
| C 983 | 16 | 1.6 | 6676 | 1 | AF031590 | AF031590 Streptomy |
| C 984 | 16 | 1.6 | 6676 | 17 | SCTRXARNP | Y16311 Streptomyce |
| C 985 | 16 | 1.6 | 6909 | 94 | AF0131816 | AF0131816 Mus muscu |
| C 986 | 16 | 1.6 | 7581 | 63 | AC014309 | AC014309 Drosophi |
| C 987 | 16 | 1.6 | 8200 | 93 | STMDNAGR | L27063 Streptomyce |
| C 988 | 16 | 1.6 | 8278 | 92 | HS285F3 | AL021457 Human DNA |
| C 989 | 16 | 1.6 | 8516 | 2 | DB8362 | DB8362 Arthropac |
| C 990 | 16 | 1.6 | 8813 | 6 | DSXDH | Y08237 D. subobscur |
| C 991 | 16 | 1.6 | 8832 | 59 | SIVRNAGENN | M97264 Shalio vir |
| C 992 | 16 | 1.6 | 9523 | 2 | AF237894SL1 | AF237894 Streptomy |
| C 993 | 16 | 1.6 | 9551 | 9 | AR076233 | AR076233 Sequence |
| C 994 | 16 | 1.6 | 9551 | 10 | I39845 | I39845 Sequence 93 |
| C 995 | 16 | 1.6 | 9551 | 97 | HUMTRHYAL | L09190 Human trich |
| C 996 | 16 | 1.6 | 9668 | 64 | AC017387 | AC017387 Lactococc |
| C 997 | 16 | 1.6 | 9910 | 1 | AE006387 | AE006387 Lactococc |
| C 998 | 16 | 1.6 | 9956 | 1 | AE000873 | AE000873 Methicoba |
| C 999 | 16 | 1.6 | 9995 | 1 | AE003891 | AE003891 Xylella f |
| C 1000 | 16 | 1.6 | 10021 | 1 | AE002537 | AE002537 Neisseria |

ALIGNMENTS-

| | | | | | | |
|------------|---|--|-------------------------|-----|-----|---------------|
| RESULT | 1 | AR047920 | 3402 bp | DNA | PAT | - 29-SEP-1999 |
| LOCUS | | Sequence 1 | from patent US 5820864. | | | |
| DEFINITION | | AR047920 | | | | |
| ACCESSION | | AR047920.1 | GI:5970263 | | | |
| VERSION | | | | | | |
| KEYWORDS | | | | | | |
| SOURCE | | Unknown. | | | | |
| ORGANISM | | Unknown. | | | | |
| REFERENCE | | 1 (bases 1 to 3402) | | | | |
| AUTHORS | | Paranios-Baccala, G., Lesenechal, M. and Jolivet, M. | | | | |
| TITLE | | Trypanosoma cruzi antigen, gene encoding therefor and methods of | | | | |
| JOURNAL | | detecting and treating chagas disease | | | | |
| FEATURES | | Patent: US 5820864-A 1 13-Oct-1998; | | | | |
| source | | Location/Qualifiers | | | | |
| BASE COUNT | | 888 a 821 c 956 g 737 t | | | | |
| ORIGIN | | /organism="unknown" | | | | |

Query Match 100.0%; Score 976; DB 9; Length 3402;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 976; Conservative 0; Mismatches 0; Indels 0;

| | | | |
|--------|--------|---|-------------|
| QY | 1 | caagfcaacgcttaacggcctttctcttcacatcgrtaacgccaagcagfagctggtcctgct | 60 |
| Db | 1232 | CAGGTACAGCCTTAACGGCTTTTGTCTTAATCGTACAGCCGAGGTAGCTGGCTCTGGCT | 1291 |
| QY | 61 | gacatgctcgaatctgttaagcaatcttcacatctccggtatccctccgcaagggaagacaag | 120 |
| Db | 1292 | GACATGTGATTTGATTGAGCATCTTCATCTCCGGTCTCCCGCAGGGAGAAACAGCAG | 1351 |
| QY | 121 | ccagagccaaaaaacaatcgtgtafctggcgagcgagcaacgagggtgtgtctcctgggaact | 180 |
| Db | 1352 | CCAGGCCAAAAAACAATCGTAGTGGCGACGCGCAAAACGGGGGTGTGTCTCCGGGCACT | 1411 |
| QY | 181 | gacgcgcgagtaagcaatcaccaataagactctctgcgcgtctgcataccctccatca | 240 |
| Db | 1412 | GACCGGGGGAATTACCAATCAATACGACTTTCGCCGCTGTGCAATCCCTTCATCA | 1471 |
| QY | 241 | ccccctgttcaagcgcagccaaagcaagccgcctcctctgcgcgagcgatcagctag | 300 |
| Db | 1472 | CCCCCTGTTTATGAGCCGCAAGCAGCAAGCGCGCCCTCTCCGCGGCGGATCGGCTAG | 1531 |
| QY | 301 | ccgaaagcgagggaagaagaatcattgtcctaactgtaatcaacgtgggattaatgtcac | 360 |
| Db | 1532 | CCGACACGGGGGACACAAGATCTATTGCTAACTTAGTAATCACTCGGGATTAAATCTACC | 1591 |
| QY | 361 | caaaagagcgtctcgcgaagcacttggagcgcgcgacagcaaggtctcaagcggttga | 420 |
| Db | 1592 | CAAAAGAGCGTCTGTCACACTGTGAAGCGCGCGCCACGACGAGGTCTAACGGCGTACGTC | 1651 |
| QY | 421 | acgactacgccccgcagcgaaacaagtcacatcagggcacaaatgtgcgaactgtga | 480 |
| Db | 1652 | ACGACTACCGCCCGCGCAGCAACAGTCCATAGGGGCAAAAGGCCGACCTGTGACGGCT | 1711 |
| QY | 481 | ggaatgtgtgagcagtaagtagtggctgcagcgcgagcctgtctcccaagcgcgagcaaa | 540 |
| Db | 1712 | GGATTTGGTGAGCTAATATGTGTTGCCAGCGCGGCTCTGTCTCCACACCGCGGCGAA | 1771 |
| QY | 541 | ccaaacaggaagaagaagacgctccgcgcgacgtgtgaagcagactccgttgcgataatgcg | 600 |
| Db | 1772 | CCAAACAGGAGAAAGAAAGGCGCTCCGCGGCAATGTGAAMCAGACTCCGTGGCGATAATCG | 1831 |
| QY | 601 | acaagcccgagcgtctcaacagcgtctctcccgcaagcgcaagagatggcgttttggcg | 660 |
| Db | 1832 | ACAGCCCGGCGCTTCAACAAGCCTCTCTCCGACAGGCCCAACGATAGCGCTTTGGCG | 1891 |
| QY | 661 | gcacgagtatcaagctcgagagcgagagttcatcagctgcgttgaagcggcgtgagatcgc | 720 |
| Db | 1892 | GCAGAGTATACAGTGTGAGGGCGAGAGTTCAATCACTGCTGTGAGCGGCTGGAGTCCGTC | 1951 |
| QY | 721 | ataaccaaacgctctcggtgtcttgaagttgtctccctgcacacatctcgaagaccatga | 780 |
| Db | 1952 | ATAACCAACAAGTCTCGGTTCTGAAGTTGTCTCCCTGACACCATTCGAAAGACCATGAA | 2011 |
| QY | 781 | caactctgaatctcgtggttttagaggcaacagatgaacagagcttcagcaagcgcgtccaca | 840 |
| Db | 2012 | CAACTCTGAATCTGGGTTTAGAGGCACAGATGACAGACTCTCAGCAGAGCCGTCCACA | 2071 |
| QY | 841 | ccgcaaaacaagccgagagacacaagctcccggaatcatccgtgttttgaagcttaccc | 900 |
| Db | 2072 | CCGCAAAACACAGCCGAAAGACACAACAAGCTCCGCAATCATCTCGTTTAGACACGTAACC | 2131 |
| QY | 901 | ctttttctaatgttgagtttccctctctcgaaatcatcgaataagaggggtgaagcgtgtgtg | 960 |
| Db | 2132 | CTTTGTTCTAATGTGGGATTCCTCTCTCTCGAACATCATCAAGAGGGGTGAAGGTGTGTG | 2191 |
| QY | 961 | aacgagagccattatgt 976 | |
| Db | 2192 | AACGAGGCCATTATGT 2207 | |
| RESULT | 2 | | |
| A48910 | A48910 | 3402 bp | DNA |
| LOCUS | | | PAT |
| | | | 07-MAR-1997 |

[illegible]

| | | | |
|------------|---|---|-------------|
| Dn | 1832 | ACAGCCGGCGCTTCACAAAGCCCTCTCTCCCGACAGCGCCACAGGATGGCGTTTGGCG | 1891 |
| Qy | 661 | gcagcagatatacagtcgagggcgaggttcacagtcgctcgagcgagtcgagtcgctc | 720 |
| Dn | 1892 | GCAGAGATATACCACTCGAGGCGAGGTCTCATGCTGCTTGAGCGGCTGAGTCCGTC | 1951 |
| Qy | 721 | ataaccaacagtcctcgaggttcgagttgctcctcgacacatcgaagagacatgaa | 780 |
| Dn | 1952 | ATAACCAACACGCTCTCGGCTTCTGAAGTTGCTCCCTGACACCATTTGAAGACCATGAA | 2011 |
| Qy | 781 | caacttcgaatctgggttttagagacacagatgacagagctcgaagagcgctccaaca | 840 |
| Dn | 2012 | CACCTTCTGAATCTGGGTTTAAAGGACACAGATGACAGAGCTGACACAGACCGCTCAACA | 2071 |
| Qy | 841 | ccgcaaacacagcgcagagacacaaagctccgcggaatcatccgltgtttgagagctcaac | 900 |
| Dn | 2072 | CCGCAACACACAGCCGACAGACACAAAGCTCCGGAATCATCTCGTTTGACAGCTACACC | 2131 |
| Qy | 901 | ctgttcctcattcgaggttcctcctcctcgcaacatcgaagggggtgagcggtggtg | 960 |
| Dn | 2132 | CTTGTCTCATTTGCGGATTCCTCTCTCGACACATCAGCAAGGGGTTGAAGCGGTGTG | 2191 |
| Qy | 961 | aacgagcgcatatgt 976 | |
| Dn | 2192 | AACGAGCGCATATGT 2207 | |
| RESULT | 3 | | |
| LOCUS | TCU24190 | 3402 bp | mRNA |
| DEFINITION | Trypanosoma cruzi Tc40 antigen (Tc40) mRNA, complete cds. | INV | 04-AUG-1997 |
| ACCESSION | U24190 | | |
| VERSION | U24190.1 | GI:790645 | |
| KEYWORDS | Trypanosoma cruzi. | | |
| SOURCE | Trypanosoma cruzi. | | |
| ORGANISM | Eukaryote; Euglenozoa; Kinetoplastida; Trypanosomatidae; | | |
| REFERENCE | 1 (bases 1 to 3402) | | |
| AUTHORS | Lesenechal, M., Duret, L., Cano, M. I., Mortara, R. A., Jolivet, M., Camargo, M. E., da Silveira, J. F., and Paranhos-Baccala, G. | | |
| TITLE | Immunodominant antigen of Trypanosoma cruzi | | |
| JOURNAL | Mol. Biochem. Parasitol. 87 (2), 193-204 (1997) | | |
| MEDLINE | 97391123 | | |
| REFERENCE | 2 (bases 1 to 3402) | | |
| AUTHORS | Lesenechal, M., Franco Da Silveira, J., Mortara, R. A., Duret, L., Camargo, M. E., Jolivet, M., and Paranhos-Baccala, G. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (06-Apr-1995) Mylene Lesenechal, BioMerieux, 22 rue Saint Jean de Dieu, Lyon 69007, France | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..3402 | | |
| | /organism="Trypanosoma cruzi" | | |
| | /strain="G" | | |
| | /db_xref="taxon:5693" | | |
| | /dev_stage="epimastigote" | | |
| | 1..35 | | |
| | /gene="Tc40" | | |
| | /note="spliced leader" | | |
| | 1..3402 | | |
| | /gene="Tc40" | | |
| | 266..3013 | | |
| | /codon_start=1 | | |
| | /product="Tc40 antigen" | | |
| | /protein_id="AAC47657.1" | | |
| | /db_xref="GI:790646" | | |
| | /translation="MTYIVDLFNHAKPSNNGRWVSDAATFNEVPEAQRVLADSQFY | | |
| | LAYTMKRHHVLRVVRNSNLKGTVAHSPKHAVKRVNVAASAGSGEFVWVY | | |
| | DETEASNGPDLAARLTKYVEKLODPYICFSPFIASERDLDLVLEFOALID | | |
| | SSSLIERDVSLEATLORNTTLRLTLPVSENSICVSGSGMPTPTPEPTMAACT | | |
| | LRNSTPSMACBEPYKALHLDLDTAEENVSLVAASTKGYOMLLTGVAEPNLKRL | | |
| gene | | | |
| CDS | | | |
| misc_RNA | | | |

| | | | | |
|---------------------------|---------------|---|---------|--------------|
| BASE COUNT | 889 a | 818 c | 958 g | 737 t |
| ORIGIN | | | | |
| Query Match | 89.5% | Score 874: | DB 96: | Length 3402: |
| Best Local Similarity | 99.8% | Pred. No. 0: | | |
| Matches 974: Conservative | 0: Mismatches | 2: Indels | 0: Gaps | 0: |
| Qy | 1 | caagtcacagcgaagcgttcttcacatcgacgcgacgagtagtcgctcgtgct | 60 | |
| Dn | 1232 | CAGGTACAGCGTAAGCGCTTTTGTCTTCAATCGTACAGCGAGAGTACGCTCGCTG | 1291 | |
| Qy | 61 | gacatgctgattcgattgaagatcttcacatcgcgtccctccgaggaagaaacag | 120 | |
| Dn | 1292 | GACATGTGGAATCGATTCACATCTTCCATCTCCGCTCCGCGAGGAGAAACAGCAG | 1351 | |
| Qy | 121 | ccagcgaacaaacacatcgtagtgcgacgcggaacacgggggtgtgtctcccgagac | 180 | |
| Dn | 1352 | CCAGGCCAAAACATTCGGTATGCGACGCGGAAACCGGGGTGTGTCTCTCGGGCACT | 1411 | |
| Qy | 181 | gacgcgcgagtagcagatcacaatcgaactctgcgcgtctgtcgaatcccttcacat | 240 | |
| Dn | 1412 | GACGCGCGAGTAGACGATCATACAAATGACTTCTGCGGTGTGATCCCTCATCA | 1471 | |
| Qy | 241 | ccccctgtttcagc | 300 | |
| Dn | 1472 | CCCCCTGTTTCAGCCGACGAGGACCGCGCTCTCTGCGCGCGCGATCGGCTGAG | 1531 | |
| Qy | 301 | ccgcgcgtggggagagaacatcgtctaatcagtagtaacagctgggagtaatgaac | 360 | |
| Dn | 1532 | CCGCACTGGGAGGACGATCTTCTATCTAGTGAATCAGCTGGGATTAATTCAC | 1591 | |
| Qy | 361 | caaaagagcgtcgtcagcagcgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc | 420 | |
| Dn | 1592 | CAAAGAGCGGTGTCAGCAGCTGAGCGCGCGGACAGAGGTTACGGCGGTGAC | 1651 | |
| Qy | 421 | acgactaccgcccgc | 480 | |
| Dn | 1652 | ACGACTACCGCCCGCCAGGAAAGTCCATACGGGCAAAAGGCGGACCTGTGACGGCT | 1711 | |
| Qy | 481 | ggattgtgtgagcagcgaatagtggtgcagcgcgcgcgcgcgcgcgcgcgcgcgcgc | 540 | |
| Dn | 1712 | GGATTGTGTGCGACTAATAGTGTGTCAGCGCGGCTGTCTCCACACCGCGCGGAAA | 1771 | |
| Qy | 541 | ccaaagagaaagaaagcgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc | 600 | |
| Dn | 1772 | CCAAAGAGGAAGAAAGGCTCCCGCGCATGTGAAGAGCTCGTGGCGATAATGCG | 1831 | |
| Qy | 601 | acagc | 660 | |
| Dn | 1832 | ACAGCGCGCGCTTCACAAAGCCCTCTCCGCGAGGCGCCAAACGATGCGCTTTTGGCG | 1891 | |
| Qy | 661 | gcagcagatatacagtcgagggcgaggttcacagtcgctcgagcgagtcgagtcgctc | 720 | |
| Dn | 1892 | GCAGAGATATACCACTCGAGGCGAGGTCTCATGCTGCTTGAGCGGCTGAGTCCGTC | 1951 | |
| Qy | 721 | ataaccaacagtcctcgaggttcgagttgctcctcgacacatcgaagagacatgaa | 780 | |
| Dn | 1952 | ATAACCAACACGCTCTCGGCTTCTGAAGTTGCTCCCTGACACCATTTGAAGACCATGAA | 2011 | |
| Qy | 781 | caacttcgaatctgggttttagagacacagatgacagagctcgaagagcgctccaaca | 840 | |

```

|||||
Db 2012 CAACCTTGATCTGGGTTTAGAGGACAGATGACAGAGCTGCAGACAGCCGTCCAAACA 2071
OY 841 ccggaacacacagccggaagacacacagctccggaatcatctcgtgtttgagacgtacacc 900
Db 2072 CCGCAACACAGACGCGAAGACACAAAGCTCCGCAATCATCCGTTTGAACACGTAAACC 2131
OY 901 ctgtctcattcgagatccctctcgcacacacacgaaggggtgaagcgtggtg 960
Db 2132 CTGTCTCATGTGGGATTCCTCTCTCGCAACATCAAGAGGGGTGAACCGTGTGTG 2191
OY 961 aacgaagccattatgt 976
Db 2192 AACGAGCGCATATATGT 2207

RESULT 4
LOCUS A48918 22 bp DNA PAT 07-MAR-1997
DEFINITION Sequence 9 from Patent WO9605312.
ACCESSION A48918
VERSION A48918.1 GI:2302577
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 22)
AUTHORS Paranhos-Baccala,G., Lesenechal,M. and Jolivet,M.
TITLE NOVEL TRYPAOSOMA CRUIZ ANTIGEN, AND GENE CODING THEREFOR; THEIR
JOURNAL APPLICATION TO THE DETECTION OF CHAGAS' DISEASE
COMMENT BIO MERIEUX (FR)
Other publication CA 2173957 960222
Other publication AU 3169195 960307
Other publication FR 2723589 960216.
FEATURES
source 1..22
/db_xref="taxon:32644"
BASE COUNT 3 a 8 c 7 g 4 t
ORIGIN
Query Match 2.3%; Score 22; DB 9; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 35 cagccgacggtagctgctcct 56
Db 1 CAGCCGACGCTAGCTGCTCCT 22

RESULT 5
LOCUS AR047927 22 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 9 from patent US 5820864.
ACCESSION AR047927
VERSION AR047927.1 GI:5970270
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 22)
AUTHORS Paranhos-Baccala,G., Lesenechal,M. and Jolivet,M.
TITLE Trypanosoma cruzi antigen, gene encoding therefor and methods of
detecting and treating chagas disease
JOURNAL Patent: US 5820864-A 9 13-OCT-1998;
FEATURES location/Qualifiers
source 1..22
/db_xref="taxon:32644"
BASE COUNT 3 a 8 c 7 g 4 t
ORIGIN

```

```

Query Match 2.3%; Score 22; DB 9; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 35 cagccgacggtagctgctcct 56
Db 1 CAGCCGACGCTAGCTGCTCCT 22

RESULT 6
LOCUS NMA42491 331801 bp DNA BCT 04-DEC-2000
DEFINITION Neisseria meningitidis serogroup A strain Z2491 complete genome;
segment 4/7
ACCESSION AL162755
VERSION AL162755.2 GI:7379742
KEYWORDS
SOURCE Neisseria meningitidis Z2491.
ORGANISM Neisseria meningitidis Z2491.
REFERENCE 1 (bases 1 to 331801)
AUTHORS Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,
Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
Davies,R.M., Davis,P., Devlin,K., Felwell,T., Hamlin,N.,
Holroyd,S., Jagsels,K., Leather,S., Moule,S., Mungall,K.,
Quail,M.A., Rajandream,M.A., Rutherford,K.M., Stammers,M.,
Skellton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.
TITLE Complete DNA sequence of a serogroup A strain of Neisseria
JOURNAL Nature 404 (6777), 502-506 (2000)
MEDLINE 20222556
REFERENCE 2 (bases 1 to 331801)
AUTHORS Parkhill,J.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT
Notes:
Details of N. meningitidis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
FEATURES
source 1..331801
/organism="Neisseria meningitidis Z2491"
/sstrain="Z2491"
/db_xref="taxon:122587"
/feature="serogroup: A"
117..1333
/gene="NMA1040"
117..1333
/gene="NMA1040"
/feature="NMA1040", possible type I restriction-modification
system specificity protein, pseudogene, len: 1217 bp;
N-termilus shows weak similarity to the C-terminal half of
TR:Q50359 (EMBL:U25415), hsdS1B, Mycoplasma pulmonis
restriction-modification enzyme subunit S1B (336 aa),
fasta scores: E(): 2.3e-06, 26.2% identity in 141 aa
overlap and to many hypothetical restriction-modification
subunits. Also similar to NMA1041, fasta scores: E():
4.8e-10, 37.0% identity in 100 aa overlap. C-terminus
similar to part of SW:U151_ECOLI (EMBL:X13145), hsdS,
Escherichia coli type I restriction enzyme EcoRI2411
specificity protein (410 aa), fasta scores: E(): 1.9e-14,
39.6% identity in 182 aa overlap. Similar to NMA1040, E():
2.9e-09, 38.0% identity in 100 aa overlap. Contains pfam
match to entry PF01420 Methylase_S, type I restriction
modification DNA specificity domain. Contains a C(8) tract
which would allow translation as an intact CDS, if
variable. Lies within a region of unusually low GC
content"
/codon_start=1

```

```
misc_feature /pseudo
/transl_table=11
/product="pseudogene (putative type I
restriction-modification system specificity protein)"
complement(157..166)
/label=DUS
/misc_feature /note="Core DNA uptake sequence: gccgcctctgaa"
171..620
/gene="NMA1040"
/note="Pfam match to entry PF01420 Methylase-S, Type I
restriction modification DNA specificity domain, score
105.50, E-value 1.1e-27"
699..706
/note="(g)8"
726..735
/gene="NMA1040"
/note="Core DNA uptake sequence: gccgcctctgaa"
/label=DUS
/misc_feature /label=DUS
740..1231
/gene="NMA1040"
/note="Pfam match to entry PF01420 Methylase-S, Type I
restriction modification DNA specificity domain, score
158.70, E-value 1e-43"
1405..1408
1415..4509
/gene="NMA1042"
1415..4509
/note="NMA1042"
/gene="NMA1042"
restriction-modification system restriction protein, len:
291 aa; similar to parts of many e.g. SW:RIR_ECOLI
(EMBL:X13145), hsdR, Escherichia coli type I restriction
enzyme Ecor124II R protein (EC 3.1.21.3) (1033 aa), fasta
scores: E(): 0.73.9% identity in 291 aa overlap (frame
1), followed by a stop codon, E(): 0.92.9% identity in
127 aa overlap (frame 1), followed by a frameshift near a
poly-A tract, E(): 0.72.5% identity in 604 aa overlap
(frame 3)"
/codon_start=1
/pseudo
/transl_table=11
/product="type I restriction-modification system
restriction protein (pseudogene)"
complement(1642..1651)
/note="Core DNA uptake sequence: gccgcctctgaa"
/label=DUS
/misc_feature /gene="NMA1042"
1687..1696
/gene="NMA1042"
/note="Core DNA uptake sequence: gccgcctctgaa"
/label=DUS
/misc_feature /label=DUS
1761..1770
/gene="NMA1042"
/note="Core DNA uptake sequence: gccgcctctgaa"
/label=DUS
/misc_feature /label=DUS
3905..3915
/gene="NMA1042"
/note="Core DNA uptake sequence: gccgcctctgaa"
/label=DUS
stem_loop /note="Stem loop containing DNA uptake sequences: aaat
gccctctgaa at ttccagacggc atttt"
4519..4528
/misc_feature /note="Core DNA uptake sequence: gccgcctctgaa"
/label=DUS
/misc_feature /label=DUS
complement(4531..4540)
/misc_feature /note="Core DNA uptake sequence: gccgcctctgaa"
/label=DUS
gene /complement(4559..6838)
/gene="c1pa"
complement(4559..6838)
CDS /gene="c1pa"
/note="NMA1045, c1pa, probable ATP-dependent protease
ATP-binding protein, len: 759 aa; similar to many e.g.
```

```
SW:CLPA_ECOLI (EMBL:M31045), c1pa, Escherichia coli
ATP-dependent C1p protease ATP-binding subunit (758 aa),
fasta scores: E(): 0.56.2% identity in 762 aa overlap.
Similar to NMA1683, fasta scores: E(): 0.37.2% identity
in 844 aa overlap. Contains Pfam match to entry PF00495
c1pa_B, Chaperonin c1pa/B, PS00871 Chaperonins c1pa/B
signature 1, PS00871 Chaperonins c1pa/B signature 2 and
two PS00017 ATP/GTP-binding site motif A (P-loop)"
/codon_start=1
/transl_table=11
/product="putative ATP-dependent protease ATP-binding
protein"
/protein_id="CAB84311.1"
/db_xref="GI:7379743"
/db_xref="SPTREMBL:Q9JVL5"
/translation="MLSPLEPHILQIOLREARNARYETISLHLLVLIEEDAAVNV
LKICGADIKVSEIOAASVANTPQIPHLIDYETQPTLGFQVIGRAMYHQSACK
AAVEPLDLIVAMSESESHAVFLKLOSTTFEVLKRICAHSPDEDESDGLGREG
EEAEQKTSLSLDYTVNLNAEVKAGRIDPLIGRKHEMERLVQILCRRRRNPLLVLEAG
VGKALALAGLAHQIVKGDIPDALKDAEYVALDMGSLAGTYRQDFEARKSVLKOLE
KIPALIFIDEIHITIGAGSTSGTGMASNLKALAKGALRCIGATYDEYRIFDK
DHLSRRFOKIDVVEPTVSEYVQILRGKPMEEGPHOYRTQALAEAAELSAARYNE
RLDPKAYDVMEDEGAQORILPKSKQKVIKQIETVIAKVARIPKTVSHDDKOYL
OFLGRDLKNMYGQENAIYALVAVKMRSGCLPDKTIGSFLESGPTGVKTEYAKO
LAYSGVPLQRFDMSEYERHVAVSKLIGAPGVGFEGGGLTEAVNQPVCVLIDE
IEKAHPDIFNVLLQVMAGKLTDNNGKSADRNVLIMTTVAGAESLSRPSLGTAKR
ERGDQMQAINKLFTPEFRNRLDALIFPAPLESEPIIVKVDVFLQLLEHRLDKVEAE
FTPALRKYSKSGFDPQMGARPMNLIOEKIRKPLADELLRGKLVGGFVRIWDMAK
EEAVLKFKSKSVKPEYEV"
complement(4637..5578)
/misc_feature /gene="c1pa"
/note="Pfam match to entry PF00495 c1pa_B, Chaperonin
c1pa/B, score 583.70, E-value 1.1e-171"
complement(5228)
/misc_feature /gene="c1pa"
/note="PS00871 Chaperonins c1pa/B signature 2"
complement(5330..5353)
/misc_feature /gene="c1pa"
/note="PS00017 ATP/GTP-binding site motif A (P-loop)"
complement(5894..5932)
/misc_feature /gene="c1pa"
/note="PS00870 Chaperonins c1pa/B signature 1"
complement(6173..6196)
/misc_feature /gene="c1pa"
/note="PS00017 ATP/GTP-binding site motif A (P-loop)"
6374..6383
/note="Core DNA uptake sequence: gccgcctctgaa"
/label=DUS
gene /complement(6842..7153)
/gene="NMA1046"
complement(6842..7153)
CDS /gene="NMA1046"
/note="NMA1046, len: 103 aa; unknown, similar to bacterial
hypothetical proteins e.g. SW:YIJA_ECOLI (EMBL:AE000100),
YIJA, Escherichia coli hypothetical protein (103 aa),
fasta scores: E(): 7.5e-14, 44.6% identity in 83 aa
overlap"
/codon_start=1
/transl_table=11
/product="hypothetical protein NMA1046"
/protein_id="CAB84312.1"
/db_xref="GI:7379744"
/db_xref="SPTREMBL:Q9JVL4"
/translation="MNHNPTDHOSEVTLSDINTOPPKRYGVFLINDYTTMFEVVEL
TEVFMIAQEOAVAVVLLVHNHGKGLCGTYTDIOTKOHOMERAKRTGHBLYKIVEE
V"
Query Match 2.3%; Score 22; DB 3; Length 331801;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 22; Mismatches 0; Indels 0; Gaps 0;
QY 419 ccagcactaccgccccgcagcg 440
|||||
```

Db 147264 CCAGACTACCGCCCGCAGCG 147285

RESULT 7

LOCUS A48919 21 bp DNA PAT 07-MAR-1997
DEFINITION Sequence 10 from Patent WO9605312.
ACCESSION A48919
VERSION A48919.1 GI:2302578
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.

REFERENCE 1 (bases 1 to 21)
AUTHORS Paranhos-Baccala,G., Lesenechal,M. and Jolivet,M.
TITLE NOVEL TRYPANOSOMA CRUZI ANTIGEN, AND GENE CODING THEREFOR; THEIR APPLICATION TO THE DETECTION OF CHAGAS' DISEASE
JOURNAL Patent: WO 9605312-A 10 22-FEB-1996;
BIO MERIEUX (FR)

COMMENT Other publication CA 2173957 960222
Other publication AU 3169195 960307
Other publication FR 2723589 960216.
Location/Qualifiers

FEATURES
source 1..21
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 6 a 7 c 3 g 5 t
ORIGIN

Query Match 2.2%: Score 21; DB 9; Length 21;
Best Local Similarity 100.0%: Pred. No. 9, 6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 956 gtgtgaacgagccattatgt 976

Db 21 GTGTGAACGAGCCATTATGT 1

RESULT 8

LOCUS A48921 21 bp DNA PAT 07-MAR-1997
DEFINITION Sequence 12 from Patent WO9605312.
ACCESSION A48921
VERSION A48921.1 GI:2302580
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.

REFERENCE 1 (bases 1 to 21)
AUTHORS Paranhos-Baccala,G., Lesenechal,M. and Jolivet,M.
TITLE NOVEL TRYPANOSOMA CRUZI ANTIGEN, AND GENE CODING THEREFOR; THEIR APPLICATION TO THE DETECTION OF CHAGAS' DISEASE
JOURNAL Patent: WO 9605312-A 12 22-FEB-1996;
BIO MERIEUX (FR)

COMMENT Other publication CA 2173957 960222
Other publication AU 3169195 960307
Other publication FR 2723589 960216.
Location/Qualifiers

FEATURES
source 1..21
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 9 a 5 c 4 g 3 t
ORIGIN

Query Match 2.2%: Score 21; DB 9; Length 21;
Best Local Similarity 100.0%: Pred. No. 9, 6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 766 cgaagagacatgaacaact 786

Db 1 CGAAGAGACCATGAACAACCT 21

RESULT 9

LOCUS AR047928 21 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 10 from patent US 5820864.
ACCESSION AR047928
VERSION AR047928.1 GI:5970271
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 21)
AUTHORS Paranhos-Baccala,G., Lesenechal,M. and Jolivet,M.
TITLE Trypanosoma cruzi antigen, gene encoding therefor and methods of detecting and treating chagas disease
JOURNAL Patent: US 5820864-A 10 13-OCT-1998;
BIO MERIEUX (FR)

FEATURES
source 1..21
/organism="unknown"

BASE COUNT 6 a 7 c 3 g 5 t
ORIGIN

Query Match 2.2%: Score 21; DB 9; Length 21;
Best Local Similarity 100.0%: Pred. No. 9, 6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 956 gtgtgaacgagccattatgt 976

Db 21 GTGTGAACGAGCCATTATGT 1

RESULT 10

LOCUS AR047930 21 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 12 from patent US 5820864.
ACCESSION AR047930
VERSION AR047930.1 GI:5970273
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 21)
AUTHORS Paranhos-Baccala,G., Lesenechal,M. and Jolivet,M.
TITLE Trypanosoma cruzi antigen, gene encoding therefor and methods of detecting and treating chagas disease
JOURNAL Patent: US 5820864-A 12 13-OCT-1998;
BIO MERIEUX (FR)

FEATURES
source 1..21
/organism="unknown"

BASE COUNT 9 a 5 c 4 g 3 t
ORIGIN

Query Match 2.2%: Score 21; DB 9; Length 21;
Best Local Similarity 100.0%: Pred. No. 9, 6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 766 cgaagagacatgaacaact 786

Db 1 CGAAGAGACCATGAACAACCT 21

RESULT 11

LOCUS AC011183 72056 bp DNA HTG 01-OCT-1999
DEFINITION Homo sapiens clone 12.B.11, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC011183
VERSION AC011183.1 GI:6006118
KEYWORDS HTG; HTGS_PHASED.
SOURCE human.
ORGANISM Homo sapiens

Query Match 2.2%: Score 21; DB 9; Length 21;
Best Local Similarity 100.0%: Pred. No. 9, 6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 766 cgaagagacatgaacaact 786

Db 1 CGAAGAGACCATGAACAACCT 21

| | | |
|-----------|--|---|
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | * |
| AUTHORS | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | * |
| TITLE | 1 (bases 1 to 72056) | * |
| REFERENCE | 1 (bases 1 to 72056) | * |
| AUTHORS | Birren,B., Linton,L., Nusbaum,C. and Lander,E. | * |
| TITLE | Homo sapiens, clone 12_B.11 | * |
| REFERENCE | Unpublished | * |
| AUTHORS | 2 (bases 1 to 72056) | * |
| TITLE | Nusbaum,C., Lander,E., Allen,N., Anderson,M., | * |
| REFERENCE | Birren,B., Linton,L., Becker,R., Boguslavsky,L., Boukhalter,B., | * |
| AUTHORS | Baldwin,J., Barna,N., Colangelo,M., Collins,S., Collamore,A., | * |
| TITLE | Brown,A., Castle,A., Colangelo,M., Collins,S., Collamore,A., | * |
| REFERENCE | Cooke,P., Dearlano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., | * |
| AUTHORS | Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., | * |
| TITLE | Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., | * |
| REFERENCE | Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., | * |
| AUTHORS | Lehoczky,J., Lieu,C., Locke,K., MacDonald,P., Marquis,N., | * |
| TITLE | McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J., | * |
| REFERENCE | Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., | * |
| AUTHORS | Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., | * |
| TITLE | Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., | * |
| REFERENCE | Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., | * |
| AUTHORS | Wyman,D., Ye,W.J., Zimmer,A. and Zody,M. | * |
| TITLE | Direct Submission | * |
| JOURNAL | Submitted (01-09-1999) Whitehead Institute/MIT Center for Genome | * |
| COMMENT | Research, 320 Charles Street, Cambridge, MA 02141, USA | * |
| | All repeats were identified using RepeatMasker: Smit, A.F.A. & | * |
| | Green, P. (1996-1997) | * |
| | http://ftp.genome.washington.edu/RM/RepeatMasker.html. | * |
| | * NOTE: This record contains 80 individual | * |
| | * sequencing reads that have not been assembled into | * |
| | * contigs. Runs of N are used to separate the reads | * |
| | * and the order in which they appear is completely | * |
| | * arbitrary. Low-pass sequence sampling is useful for | * |
| | * identifying clones that may be gene-rich and allows | * |
| | * overlap relationships among clones to be deduced. | * |
| | * However, it should not be assumed that this clone | * |
| | * will be sequenced to completion. In the event that | * |
| | * the record is updated, the accession number will | * |
| | * be preserved. | * |
| | 1 | * |
| | 898: contig of 898 bp in length | * |
| | gap of unknown length | * |
| | 899 | * |
| | 1785: contig of 887 bp in length | * |
| | gap of unknown length | * |
| | 1786 | * |
| | 2678: contig of 893 bp in length | * |
| | gap of unknown length | * |
| | 2679 | * |
| | 3566: contig of 888 bp in length | * |
| | gap of unknown length | * |
| | 3567 | * |
| | 4474: contig of 908 bp in length | * |
| | gap of unknown length | * |
| | 4475 | * |
| | 5394: contig of 920 bp in length | * |
| | gap of unknown length | * |
| | 5395 | * |
| | 6335: contig of 941 bp in length | * |
| | gap of unknown length | * |
| | 6336 | * |
| | 7293: contig of 958 bp in length | * |
| | gap of unknown length | * |
| | 7294 | * |
| | 8172: contig of 878 bp in length | * |
| | gap of unknown length | * |
| | 8173 | * |
| | 9058: contig of 886 bp in length | * |
| | gap of unknown length | * |
| | 9059 | * |
| | 9962: contig of 904 bp in length | * |
| | gap of unknown length | * |
| | 9963 | * |
| | 10708: contig of 746 bp in length | * |
| | gap of unknown length | * |
| | 10709 | * |
| | 11617: contig of 909 bp in length | * |
| | gap of unknown length | * |
| | 11618 | * |
| | 12521: contig of 904 bp in length | * |
| | gap of unknown length | * |
| | 12522 | * |
| | 13412: contig of 891 bp in length | * |
| | gap of unknown length | * |
| | 13413 | * |
| | 14301: contig of 889 bp in length | * |
| | gap of unknown length | * |
| | 14302 | * |
| | 15220: contig of 919 bp in length | * |
| | gap of unknown length | * |
| | 15221 | * |
| | 16096: contig of 876 bp in length | * |
| | gap of unknown length | * |
| | 16097 | * |
| | 17038: contig of 942 bp in length | * |
| | gap of unknown length | * |
| | 17039 | * |
| | 17964: contig of 926 bp in length | * |
| | gap of unknown length | * |
| | 17965 | * |
| | 18911: contig of 947 bp in length | * |
| | gap of unknown length | * |
| | 18912 | * |
| | 19842: contig of 931 bp in length | * |
| | gap of unknown length | * |
| | 19843 | * |
| | 20736: contig of 894 bp in length | * |
| | gap of unknown length | * |
| | 20737 | * |
| | 21598: contig of 862 bp in length | * |
| | gap of unknown length | * |
| | 21599 | * |
| | 22508: contig of 910 bp in length | * |
| | gap of unknown length | * |
| | 22509 | * |
| | 23395: contig of 887 bp in length | * |
| | gap of unknown length | * |
| | 23396 | * |
| | 24301: contig of 906 bp in length | * |
| | gap of unknown length | * |
| | 24302 | * |
| | 25201: contig of 900 bp in length | * |
| | gap of unknown length | * |
| | 25202 | * |
| | 26120: contig of 919 bp in length | * |
| | gap of unknown length | * |
| | 26121 | * |
| | 27005: contig of 885 bp in length | * |
| | gap of unknown length | * |
| | 27006 | * |
| | 27865: contig of 860 bp in length | * |
| | gap of unknown length | * |
| | 27866 | * |
| | 28756: contig of 891 bp in length | * |
| | gap of unknown length | * |
| | 28757 | * |
| | 29667: contig of 911 bp in length | * |
| | gap of unknown length | * |
| | 29668 | * |
| | 30545: contig of 878 bp in length | * |
| | gap of unknown length | * |
| | 30546 | * |
| | 31489: contig of 944 bp in length | * |
| | gap of unknown length | * |
| | 31490 | * |
| | 32445: contig of 956 bp in length | * |
| | gap of unknown length | * |
| | 32446 | * |
| | 33362: contig of 917 bp in length | * |
| | gap of unknown length | * |
| | 33363 | * |
| | 34288: contig of 926 bp in length | * |
| | gap of unknown length | * |
| | 34289 | * |
| | 35195: contig of 907 bp in length | * |
| | gap of unknown length | * |
| | 35196 | * |
| | 36090: contig of 895 bp in length | * |
| | gap of unknown length | * |
| | 36091 | * |
| | 36976: contig of 886 bp in length | * |
| | gap of unknown length | * |
| | 36977 | * |
| | 37702: contig of 726 bp in length | * |
| | gap of unknown length | * |
| | 37703 | * |
| | 38398: contig of 636 bp in length | * |
| | gap of unknown length | * |
| | 38399 | * |
| | 39314: contig of 916 bp in length | * |
| | gap of unknown length | * |
| | 39315 | * |
| | 40226: contig of 912 bp in length | * |
| | gap of unknown length | * |
| | 40227 | * |
| | 41133: contig of 907 bp in length | * |
| | gap of unknown length | * |
| | 41134 | * |
| | 42015: contig of 882 bp in length | * |
| | gap of unknown length | * |
| | 42016 | * |
| | 42879: contig of 864 bp in length | * |
| | gap of unknown length | * |
| | 42880 | * |
| | 43787: contig of 908 bp in length | * |
| | gap of unknown length | * |
| | 43788 | * |
| | 44732: contig of 945 bp in length | * |
| | gap of unknown length | * |
| | 44733 | * |
| | 45645: contig of 913 bp in length | * |
| | gap of unknown length | * |
| | 45646 | * |
| | 46606: contig of 961 bp in length | * |
| | gap of unknown length | * |
| | 46607 | * |
| | 47527: contig of 921 bp in length | * |
| | gap of unknown length | * |
| | 47528 | * |
| | 48457: contig of 930 bp in length | * |
| | gap of unknown length | * |

```

* 48458 49401: contig of 944 bp in length
* gap of unknown length
* 49402 50300: contig of 899 bp in length
* gap of unknown length
* 50301 51179: contig of 879 bp in length
* gap of unknown length
* 51180 52130: contig of 951 bp in length
* gap of unknown length
* 52131 53074: contig of 944 bp in length
* gap of unknown length
* 53075 54010: contig of 936 bp in length
* gap of unknown length
* 54011 54906: contig of 896 bp in length
* gap of unknown length
* 54907 55797: contig of 891 bp in length
* gap of unknown length
* 55798 56687: contig of 890 bp in length
* gap of unknown length
* 56688 57579: contig of 892 bp in length
* gap of unknown length
* 57580 58443: contig of 864 bp in length
* gap of unknown length
* 58444 59345: contig of 902 bp in length
* gap of unknown length
* 59346 60251: contig of 906 bp in length
* gap of unknown length
* 60252 61197: contig of 946 bp in length
* gap of unknown length
* 61198 62066: contig of 869 bp in length
* gap of unknown length
* 62067 62960: contig of 894 bp in length
* gap of unknown length
* 62961 63871: contig of 911 bp in length
* gap of unknown length
* 63872 64769: contig of 898 bp in length
* gap of unknown length
* 64770 65714: contig of 945 bp in length
* gap of unknown length
* 65715 66596: contig of 882 bp in length
* gap of unknown length
* 66597 67499: contig of 903 bp in length
* gap of unknown length
* 67500 68408: contig of 909 bp in length
* gap of unknown length
* 68409 69331: contig of 923 bp in length
* gap of unknown length
* 69332 70264: contig of 933 bp in length

```

Query Match 2.0%; Score 20; DB 62; Length 72056;
 Best local Similarity 100.0%; Pred. No. 10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 538 aaacacacgaggaagaaaa 557
 Db 39898 AACCAACAGAGAGAGAAA 39917

RESULT 12
 AC011438/C AC011438 91470 bp DNA PLN 11-OCT-2000
 DEFINITION Genomic sequence for Arabidopsis thaliana BAC T23G18 from
 chromosome I, complete sequence.
 ACCESSION AC011438 GI:6289106
 VERSION AC011438.3
 KEYWORDS HTG.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 91470)
 Shin,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C.,
 Khan,S., Kim,C., Altafi,H., Bel,Q., Chin,C., Chlou,J., Choi,E.,
 Khan,S., Kim,C., Altafi,H., Bel,Q., Chin,C., Chlou,J., Choi,E.,

Comm,L., Conway,A., Gonzales,A., Hansen,N., Howing,B., Koo,T.,
 Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S.,
 Mukharisky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J.,
 Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G.,
 Federspiel,N.A., Theologis,A. and Ecker,J.R.
 Genomic sequence for Arabidopsis thaliana BAC T23G18 from
 chromosome I
 Unpublished
 2< (bases 1 to 91470)
 Ecker,J.R.
 Direct Submission
 Submitted (07-OCT-1999) Arabidopsis thaliana Genome Center,
 Department of Biology, University of Pennsylvania, 38th Street and
 Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
 3 (bases 1 to 91470)
 Ecker,J.R.
 Direct Submission
 Submitted (09-NOV-1999) Arabidopsis thaliana Genome Center,
 Department of Biology, University of Pennsylvania, 38th Street and
 Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
 4 (bases 1 to 91470)
 Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C.,
 Shin,P., Altafi,H., Bel,Q., Chin,C., Chlou,J., Choi,E., Conn,L.,
 Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B.,
 Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharisky,N.,
 Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A.,
 Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Davis,R.,
 Federspiel,N., Theologis,A. and Ecker,J.
 Direct Submission
 Submitted (15-DEC-1999) Arabidopsis thaliana Genome Center,
 Department of Biology, University of Pennsylvania, 38th and
 Hamilton Walk, Philadelphia, PA 19104-6018, USA
 5 (bases 1 to 91470)
 Cheuk,R., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C.,
 Khan,S., Kim,C., Shin,P., Altafi,H., Bel,Q., Chin,C., Chlou,J.,
 Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B.,
 Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S.,
 Mukharisky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J.,
 Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Davis,R.,
 Federspiel,N., Theologis,A. and Ecker,J.
 Direct Submission
 Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center,
 Department of Biology, University of Pennsylvania, 38th and
 Hamilton Walk, Philadelphia, PA 19104-6018, USA
 6 (bases 1 to 91470)
 Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C.,
 Shin,P., Altafi,H., Bel,Q., Chin,C., Chlou,J., Choi,E., Conn,L.,
 Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B.,
 Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharisky,N.,
 Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A.,
 Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Davis,R.,
 Federspiel,N., Theologis,A. and Ecker,J.
 Direct Submission
 Submitted (11-OCT-2000) Arabidopsis thaliana Genome Center,
 Department of Biology, University of Pennsylvania, 38th and
 Hamilton Walk, Philadelphia, PA 19104-6018, USA
 On Nov 9, 1999, this sequence version replaced gi:6139057.
 Location/Qualifiers
 1. 91470
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="T23G18"
 /complement(join(831..899,1015..1092,1195..1329,1360..1473,
 1513..1586,1687..1758,1838..1899,2663..2775,2882..3013,
 3472..3559,3638..3732,3827..4057,4288..4404,4481..4582,
 4711..4827,4948..5733))
 /note="putative DNA ligase sp|Q42572"
 /codon_start=1
 /evidence="not_experimental"
 /product="T23G18.1"
 /protein_id="AAF18258.1"
 /db_xref="GI:6579215"

CDS

COMMENT
 FEATURES

source

/translatabion="MLAIRSNMNAKRCITPSTLCRTQISOPSSVYLISFESROISHLIRLHNS
 HRAMSSPSPADMLASNAARAUKKRCITPSTLCRTQISOPSSVYLISFESROISHLIRLHNS
 TLPRKEDLLEPVSDSANPSRSDTSSIAEDSKTAKKAKTISKTDMDKSTIGLCKKPPND
 FDPKMSCKMEKERPELFEVLAFLDILISENGEITVTDLCNMLRYLIATPEEDVATV
 VLSANETAPAHGEVLELIGESTETIKAIASEBGTIDBYHCKOITLEDLIDLVAGKSSV
 TOTAMEKEPPELVAVDFDRIOIAKSGSDSPNKPKKRLKMAALIVATQEDCPYLITRLL
 OAKLIGESGQVLAALGOAAVYNEHSPKRPKPKKRLKMAALIVATQEDCPYLITRLL
 VALLSGSGLVLELIGESTETIKAIASEBGTIDBYHCKOITLEDLIDLVAGKSSV
 OIHMEGDCFEELYSYNAEENRCKPQVVALSLPKRKSYSKSLDCEVAPFDEKPKKI
 LPOLISTRANKNVYNDIKVYCFILAFDMLYINGCOLIOENLIRKREKILSEEDP
 IDVDEPVSATSDIDEIOFLDASVDVIGDSDVILAFEBHGRKGTGYGAFLVAG
 GFDVKEEIOSTICKITGEFDMALDENSSLSRQVLAIPKIEFLPENCDSNIQOYRRLAC
 DLSNDEPPEPELVAVDFDRIOIAKSGSDSPNKPKKRLKMAALIVATQEDCPYLITRLL
 IYDVPKGSILPEPRLLRKISIRKIVWEKKKQVLELKWENVAADLTISPVRNAKGT
 complement:join(6737.7168.7256.7909.8003.9532
 10198.10618.10842.11504.11589.12614.12710.12982,
 14353.14725.14823.15484.15389.16608.16666.17013,
 17659.18255))
 /note="putative Na/H antiporter g114063740; similar to
 ESTs g1141634, g114A586073"
 /codon_start=1
 /evidence=not experimental
 /product="733618.2
 /protein_id="AA18257.1"
 /db_xref="GI:6579214"
 /translatabion="AAPPBAPADQAPARTOPOLQPOSQAPAPLGRMRNMLITVAL
 VLGLVAGMLATITVALTPKRLIYVEAASVQEFALIGNDDHINKAFSVIKSNPEE
 KHSVRSYSHSTLHHNNOVAHKNTSPKORKNETRIETOLVSHVSLKSNAPLE
 RAERKSGITEMEVYITATRSYKTYFRSRRKNTIACTCPKKMLKESNENAEAGY
 NATIPGDSYHETGSEDSKCEQPIILVNSGVCBKLDPIRKMTIDYLSPEDES
 VIVLCLMOPFYSLSKILGLPVPKITSMTAVRTSSSESGAALSTONLPLNPTIOH
 ILFPDTRKREYELTGEMFVRLFEQVAGMDWVRKGTQVILVGEIARVILPIIAA
 NMVEKLEKTEGKVLGTGMFVRLFEQVAGMDWVRKGTQVILVGEIARVILPIIAA
 VADGEGFVNLELVAAMDVRPVALOGVIGIVIMVWVRAMEVIRKTPREVRV
 KECFVILIIAAGFVLEKELHMPAPBELLGVCVHPGRPSOLVCEENSGNIGI
 LLPLPEFMSLODQBPVALNGOLPHRPGQLEATITVYVAKITFEMIPALAK
 MLPDSFPMALILNSKGIYELCELEGVSNVLAHSTTMMATVILSTISYLLIR
 LDDSKRTISPKRNLMSKILSDSEKFLVLCIKHADHISGIMTALLOSFLHESTISCY
 PRVHVEGLNDPRLISHOMKAPENRYSNNVLAIDNFKRWKSTISLEFPLSTISCY
 PRVHVEGLNDPRLISHOMKAPENRYSNNVLAIDNFKRWKSTISLEFPLSTISCY
 IYVHMOEYSLALDQASFTLMPHITLMSLDQTVSDVDYRRNRNALVLRQASVYG
 IFVHMOEYSLALDQASFTLMPHITLMSLDQTVSDVDYRRNRNALVLRQASVYG
 DFTGKMOOMLDQASFTLMPHITLMSLDQTVSDVDYRRNRNALVLRQASVYG
 LEVGRSNGYGVDSVLSAEMRPELDELITSEFFTLRWLIDSPFKIGLAVPRRTYMM
 VNSHIEYMGVTRKCMNPEYELPDELITSEFFTLRWLIDSPFKIGLAVPRRTYMM
 IVRLIGTGCHFSNKSMDHIDIFFPDNRKVAETLDAFGVLVFLWLGKVMDELPRK
 TERKSSVYELVILVILPGSLTEFRERKRGDSILMEERLILFLOSISAFPSIDPL
 LKDOIKISEFERVILVILPGSLTEFRERKRGDSILMEERLILFLOSISAFPSIDPL
 VVRPMYAWIKOTPEGRPKDYLSISFGLAFCEFFENKVIHLFGASGFVGLVAV
 NGYPLGTUILLIOKESFENSGISITPFSILTMOMVDILRFKESDILRMEGOIEVISE
 TLLVNTFVYVYITAAVAKMLROSEFALVLSNKGIFELAYVAAVEAKILPEVAF
 TILAYVILNLSFFIMLELVDHPYTRFRCYKRRNIGILKQGAALOCINCYKPRDITTT
 SMTDLLETFSPDSPMACNLIHLELVQDANRFTSHQLOKREPGSTISDNYISF
 GEGFOKOFPEYSLDTIFSVYSVOHMEIDICMLASLSLITVLPFHRTWSDRSTVYS
 IDDDNLRQAFENYSLDTIFSVYSVOHMEIDICMLASLSLITVLPFHRTWSDRSTVYS
 IYRMARTEKRTLITIFRPEKSEMDDEQOOSILNKESVTIGVINKENAKVARY
 IDKASDSESTRILIRAMANDVDFVSGSGSGTGATGTSISMTENELSGDGLIAA
 SHEVSKSTLMOCLETRVLIIVLITLASSHOCPEFLPIYLRDOLSDMDANATW
 RKEPMMNDRKDAEMGTAKCFVDSPIHMLASAKEMASGSGMDPEYELPDELIT
 LLSTFLMLRPEMLRKGLVPLPKFTSSMLVVRNGVOISACHNGSNHTPLPLFLK
 GAVGLSEMGSGMOLCKRIELHYNOYMRVPTIOTIAFAVLTWTFPMGMPDVGIMKK
 SGTVSGVIGTISMIIPMOGLKILSRKESKSLITJMEETVMTFPMGMPDVGIMKK
 LDKLIVHTDFQIOAKAMNDILAFVLSVYSVSDRQEGOGAGLAFNAFIFVYL
 VROEMALVIRHPEAGPQAVNYVILGLLAATVLSYMSRSLFPGGCAFLGALVAPG
 PGLSGVFOKQPSRNDGILFLPGLSMILKIDMSLPARKENGRHLGHNYBCSPPLP
 IVYIAKTRFSLAALAPVILPDSITILGVIYKSSSELDVILTAPEKDIISLEVL
 LGVYILVNSLFLAPVIAHFLYDSSKRVCGGRMLKRPENQVILAKNPNTISML
 LRAVSPDSMEPCVCLHILIELGQATPFIISHQLOPKKGSYSSENYISLQFOD
 IYVSPASISNMFSTISAKMHQOICWFLASOGSNLILSHPRMEGANYISODPOT
 RSLNLTIVLRKACSGVIFVYRKPIDOTKLESCVRLIYVGGNDKEALALADHMG
 NQOQVITVLRKILPVSYGADSSLRTHSQMOTMDMNEEDQROBKDTIIDMYQDETYSK
 TLHSVSYDIDLEIVGRSGSVGTTVTRGLDMDMEBELGVLGDLIASPPSRASVAVV
 000E"
 complement(20859.21590)

| Query Match | Best Local Similarity | 2.0%: Score 20: DB 12: Length 91470: |
|---|--|--------------------------------------|
| Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | |
| Qy 212 ctctctgcgcctgctgcatcc 231 | | |
| DB 42747 CTTCTGCGCGTCTGCATCC 42728 | | |
| | | |
| RESULT 13 | | |
| AL157771 | 115243 bp | PRI 28-Jan-2001 |
| LOCUS | Human DNA sequence from clone RP11-67J18 on chromosome 13, complete | |
| DEFINITION | sequence. | |
| ACCESSION | AL157771 | |
| VERSION | AL157771.11 | GI:12597045 |
| KEYWORDS | HTG. | |
| SOURCE | human. | |
| ORGANISM | Homo sapiens | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| AUTHORS | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | |
| TITLE | 1 (bases 1 to 115243) | |
| JOURNAL | Oliver, K. | |
| COMMENT | Direct Submission Submitted (28-Jan-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk requests: clonerquests@sanger.ac.uk On Jan 29, 2001 this sequence version replaced gi:12578162. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission | |

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/Projects/C.elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr13> RP11-67J18 is from the library RPc1-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/Bacpac/home.htm> VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-67J18. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP11-67J18 is at 1 in this sequence. The true left end of clone RP11-141M24 is at 115144 in this sequence. The true right end of clone RP11-465C24 is at 93228 in this sequence.

FEATURES

source

1. Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone_11b="RPc1-11.1"
6..140
/note="MUT1A2 repeat: matches 239. .374 of consensus"
repeat_region
1758..2128
/note="THE1C repeat: matches 1. .371 of consensus"
repeat_region
2649..2734
/note="L2 repeat: matches 2575. .2670 of consensus"
repeat_region
3230..3626
/note="THE1C repeat: matches 1. .371 of consensus"
repeat_region
3627..5155
/note="THE1C-internal repeat: matches 1. .1580 of consensus"
repeat_region
5158..5525
/note="THE1C repeat: matches 3. .371 of consensus"
repeat_region
6411..6476
/note="33 copies 2 mer ta 66% conserved"
repeat_region
6520..6804
/note="AluX repeat: matches 7. .292 of consensus"
repeat_region
7211..7363
/note="L2 repeat: matches 2547. .2750 of consensus"
repeat_region
7923..8133
/note="MIR repeat: matches 31. .237 of consensus"
repeat_region
9831..9899
/note="AluY repeat: matches 241. .309 of consensus"
repeat_region
10360..10555
/note="L2 repeat: matches 2509. .2740 of consensus"
repeat_region
11346..13655
/note="L1M87 repeat: matches 3603. .5986 of consensus"
repeat_region
14729..14939
/note="L2 repeat: matches 2112. .2336 of consensus"
repeat_region
15021..15202
/note="MER3A repeat: matches 3. .189 of consensus"
repeat_region
15405..15503
/note="L2 repeat: matches 2626. .2732 of consensus"
repeat_region
15753..16043
/note="L2 repeat: matches 1808. .2126 of consensus"
repeat_region
16705..16936
/note="LIME2 repeat: matches 5910. .6148 of consensus"
repeat_region
17834..18076
/note="MIR repeat: matches 4. .258 of consensus"

repeat_region
19095..19160
/note="L2 repeat: matches 2634. .2701 of consensus"
repeat_region
19187..19455
/note="MERR44A repeat: matches 501. .331 of consensus"
repeat_region
19653..19933
/note="MERR44A repeat: matches 1. .258 of consensus"
repeat_region
20608..20816
/note="L1M87 repeat: matches 174. .389 of consensus"
repeat_region
22649..22692
/note="L2 repeat: matches 299. .1401 of consensus"
repeat_region
24393..24698
/note="AluX repeat: matches 1. .309 of consensus"
repeat_region
24699..24730
/note="L6 copies 2 mer ta 84% conserved"
repeat_region
24713..24831
/note="7 copies 17 mer 72% conserved"
repeat_region
24799..24856
/note="29 copies 2 mer at 70% conserved"
repeat_region
26293..26446
/note="MIR repeat: matches 40. .206 of consensus"
repeat_region
27366..27424
/note="L2 repeat: matches 598. .660 of consensus"
repeat_region
29663..29966
/note="AluSq repeat: matches 1. .304 of consensus"
repeat_region
31046..31563
/note="MUT2B repeat: matches 1. .448 of consensus"
repeat_region
31612..32579
/note="L2 repeat: matches 299. .1401 of consensus"
repeat_region
32652..32780
/note="L2 repeat: matches 1577. .1709 of consensus"
repeat_region
32994..33118
/note="L1M2 repeat: matches 5498. .5625 of consensus"
repeat_region
33371..33666
/note="AluX repeat: matches 1. .299 of consensus"
repeat_region
34237..34272
/note="18 copies 2 mer tg 100% conserved"
repeat_region
34274..34544
/note="AluSc repeat: matches 1. .271 of consensus"
repeat_region
35460..35724
/note="AluSp repeat: matches 1. .265 of consensus"
repeat_region
35764..36107
/note="L2 repeat: matches 2349. .2688 of consensus"
repeat_region
36950..37184
/note="L2 repeat: matches 1937. .2198 of consensus"
repeat_region
37349..37413
/note="L2 repeat: matches 1759. .1824 of consensus"
repeat_region
37430..37833
/note="L2 repeat: matches 1837. .2255 of consensus"
repeat_region
38102..38355
/note="L2 repeat: matches 1171. .1445 of consensus"
repeat_region
39865..40002
/note="MIR repeat: matches 57. .187 of consensus"
repeat_region
40308..40584
/note="AluX repeat: matches 1. .302 of consensus"
repeat_region
40875..41431
/note="L1M21 repeat: matches 5588. .6157 of consensus"
repeat_region
41431..41691
/note="L1M22 repeat: matches 5639. .5892 of consensus"
repeat_region
42213..42280
/note="34 copies 2 mer ct 75% conserved"
repeat_region
46770..46819
/note="25 copies 2 mer tc 86% conserved"
repeat_region
46907..47093
/note="L1P4 repeat: matches 5958. .6144 of consensus"
repeat_region
47194..47232
/note="L1P3 repeat: matches 502. .537 of consensus"
repeat_region
48816..48996
/note="MIR repeat: matches 8. .231 of consensus"
repeat_region
50009..50070
/note="31 copies 2 mer tg 80% conserved"
repeat_region
50276..51464
/note="L1M2 repeat: matches 4936. .6155 of consensus"
repeat_region
51642..51691

```

repeat_region /note="25 copies 2 mer at 86% conserved"
52061..52411
/note="THE1A repeat: matches 1..354 of consensus"
52412..53977
/note="THE1A-internal repeat: matches 1..1580 of
consensus"
53978..54330
/note="THE1A repeat: matches 1..354 of consensus"
55056..55188
/note="LIMB2 repeat: matches 4769..4900 of consensus"
55834..55990
/note="LIP repeat: matches 3602..3764 of consensus"
56216..56803
/note="LIM4 repeat: matches 2395..3180 of consensus"
57052..57093
/note="21 copies 2 mer tg 78% conserved"
57181..57250
/note="35 copies 2 mer at 68% conserved"
57319..57388
/note="L2 repeat: matches 2035..2104 of consensus"
57655..57720
/note="L2 repeat: matches 2635..2701 of consensus"
57823..57981
/note="FRAM repeat: matches 1..159 of consensus"
58139..60124
/note="LIPBA repeat: matches -279..2655 of consensus"
60112..61146
/note="LIPBA repeat: matches -1545..-412 of consensus"
62246..62346
/note="MIR repeat: matches 81..187 of consensus"
62381..63266
/note="Alu10 repeat: matches 13..272 of consensus"
64366..64467

```

```

Query Match 2.0% Score 20: DB 90: Length 115243;
Best Local Similarity 100.0%; Pred. NO. 9.7;
Matches 20: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 719 tcatatccacacgtctcgcg 738
|||||
Db 76310 TCATATACCAACACGCTCTCG 76329

```

```

RESULT 14
AC087859/c 153259 bp DNA PRI 02-FEB-2001
LOCUS Homo sapiens chromosome 3 clone RP11-34L16 map 3p, complete
DEFINITION sequence.
ACCESSION AC087859
VERSION AC087859.1 GI:12656789
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 153259)
Zhang,L., Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C.,
Chen,J., Ding,H., Dong,W., Fan,H., Feng,X., Jin,Y., Kang,N.,
Gu,X., Guo,D., Guo,Z., He,L., Hu,S., Huang,F., Jin,Y., Kang,N.,
Li,C., Li,C., Li,F., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y.,
Liu,N., Liu,B., Liu,Y., Li,W., Li,W., Li,Y., Luo,C., Luo,J.,
Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y.,
Tan,X., Tao,R., Wang,H., Wang,J., Wang,J., Wang,L.,
Wang,L., Wang,R., Wang,X., Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F.,
Xuan,Z., Xue,Y., Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G.,
Zhang,H., Zhang,H., Zhang,M., Zhang,X., Zhang,Y., Zhang,Y.,
Zhang,Y., Zhang,Z., Zhu,B., Zhu,N., Yu,J. and Yang,H.
Chromosome 3p genomic sequence
Unpublished
TITLE JOURNAL
REFERENCE 2 (bases 1 to 153259)
AUTHORS Chen,L., Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C.,
Chen,J., Ding,H., Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q.,
Gu,X., Guo,D., Guo,Z., He,L., Hu,S., Huang,F., Jin,Y., Kang,N.,

```

```

Li,C., Li,C., Li,F., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y.,
Liu,N., Liu,B., Liu,Y., Li,W., Li,W., Li,Y., Luo,C., Luo,J.,
Liu,Y., Qi,Q., Qi,X., Song,L., Song,S., Sun,M., Sun,W., Sun,Y.,
Tan,X., Tao,R., Wang,H., Wang,J., Wang,J., Wang,L.,
Wang,L., Wang,R., Wang,X., Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F.,
Xuan,Z., Xue,Y., Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G.,
Zhang,H., Zhang,H., Zhang,M., Zhang,X., Zhang,X., Zhang,Y.,
Zhang,Y., Zhang,Z., Zhu,B., Zhu,N., Yu,J. and Yang,H.
Direct Submission
Submitted (02-FEB-2001) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China
-----Genome Center
Center:Beijing Center
Center code:Beijing
Website:http://hgsc.igtp.ac.cn
http://www.genomics.org.cn
Contact:hgsc@igtp.ac.cn
----- Project Information
Center project name:1% project
Center clone name: RP11-34L16
----- Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator; ET 55% of reads
Chemistry: Dye-terminator Big Dye; 45% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 952 bases at least Q40
Consensus quality: 1082 bases at least Q30
Insert size: 1160; sum-of-contigs
Quality coverage: 1.59x in Q20 bases;sum-of-contigs
-----

```

```

FEATURES
source
1..153259
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"
/clone="RP11-34L16"
BASE COUNT 41764 a 33139 c 34151 g 44205 t
ORIGIN

```

```

Query Match 2.0% Score 20: DB 88: Length 153259;
Best Local Similarity 100.0%; Pred. NO. 9.3;
Matches 20: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 813 gacagagctgcagcagacc 832
|||||
Db 3666 GACAGAGCTGCAGCAGACC 3647

```

```

RESULT 15
AC018355/c 155075 bp DNA HTG 07-SEP-2000
LOCUS Homo sapiens chromosome 3 clone RP11-34L16, WORKING DRAFT SEQUENCE,
DEFINITION 12 unordered pieces.
ACCESSION AC018355
VERSION AC018355.14 GI:9966753
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 155075)
Muzny,D.M., Adams,C., Bailey,M., Barbarta,J., Blankenburg,K.,
Booth,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunc,C.,
Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
David,R., Delgado,O., Deshazo,D., Ding,Y., Donah-Rashid,N.,
Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
Forcum-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hughes,M.,
Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,

```

Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z., Licharge,O., Liu,J., Liu,W., Logan,O., Lozano,R.J., Lu,J., Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M., Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S., Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L., Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S., Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sucgang,R., Taber,P., Taylor,T., Vasquez,L., Vinson,R., Vo,O., Wahab,M., Watlington,S., Weinstein,G., Weinstein,I.R., Williamson,A., Worley,K., Wren,J., Wrenford,G., Yu,W., Zhou,X., Nelson,D. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 155075)
Worley,K.C.

Direct Submission
Submitted (10-DEC-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Sep 5, 2000 this sequence version replaced gi:8493479.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: HMYJ
Center clone name: RP11-34L6
----- Summary Statistics -----
Sequencing vector: M13: L08821
Assembly program: Phrap: version 0.990329
Consensus quality: 138685 bases at least Q40
Consensus quality: 147372 bases at least Q30
Consensus quality: 149805 bases at least Q20
Estimated insert size: 150288; sum-of-contigs estimation
Estimated insert size: 161640; agarose-fp estimation
Quality coverage: 4.3x in Q20 bases; agarose-fp estimation
Quality coverage: 4.6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 48839: contig of 48839 bp in length
* 48840 48839: gap of unknown length
* 48940 81367: contig of 32428 bp in length
* 81368 81467: gap of unknown length
* 81468 93764: contig of 12297 bp in length
* 93765 93864: gap of unknown length
* 93865 107706: contig of 13842 bp in length
* 107707 107806: gap of unknown length
* 107807 120074: contig of 12268 bp in length
* 120075 120174: gap of unknown length
* 120175 130823: contig of 10649 bp in length
* 130824 130923: gap of unknown length
* 130924 137196: contig of 6273 bp in length
* 137197 137296: gap of unknown length
* 137297 143104: contig of 5808 bp in length
* 143105 143204: gap of unknown length
* 143205 146435: contig of 3231 bp in length
* 146436 146535: gap of unknown length
* 146536 149651: contig of 3116 bp in length
* 149652 149751: gap of unknown length
* 149752 153787: contig of 4036 bp in length
* 153788 153887: gap of unknown length
* 153888 155075: contig of 1188 bp in length.

FEATURES
source Location/Qualifiers
1..155075
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-34L6"

BASE COUNT 42142 a 33789 c 33967 g 44018 t 1159 others
ORIGIN

Query Match: 2.0%; Score 20; DB 65; Length 155075;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 813 gacagagctcagcagagcc 832
|||||
Db 134584 GACGAGCTCAGCAGAGCC 134565

RESULT 16
AC034187/c 171714 bp DNA PRI 09-OCT-2000
LOCUS Homo sapiens chromosome 3 clone RP11-202A21 map 3p, complete
DEFINITION sequence.
ACCESSION AC034187
VERSION AC034187.4 GI:10719838
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 171714)
Zhang,L., Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C.,
Chen,J., Ding,H., Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q.,
Gu,X., Guo,D., He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C.,
Li,C., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B.,
Liu,Y., Li,W., Li,W., Li,Y., Luo,J., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,L.,
Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J.,
Wang,J., Wang,L., Wang,L., Wang,R., Wang,X., Wang,X., Wang,Y.,
Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X., Yu,B.,
Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N.,
Yu,J. and Yang,H.

Chromosome 3p genomic sequence
Unpublished
2 (bases 1 to 171714)
Li,G., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B., Bao,W.,
Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y., Qi,X.,
Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y., Li,G.,
Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D.,
Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
and Yang,H.

Direct Submission
Submitted (05-APR-2000) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China
3 (bases 1 to 171714)
Chen,J., Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C.,
Zhang,L., Ding,H., Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q.,
Gu,X., Guo,D., He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C.,
Li,C., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B.,
Liu,Y., Li,W., Li,W., Li,Y., Luo,J., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,L.,
Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J.,
Wang,J., Wang,L., Wang,L., Wang,R., Wang,X., Wang,X., Wang,Y.,
Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X., Yu,B.,
Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N., Yu,J. and
Yang,H.

Direct Submission
Submitted (09-OCT-2000) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China
On Oct 9, 2000 this sequence version replaced gi:9858933.

TITLE
JOURNAL
AUTHORS
COMMENT

-----Genome Center
Center:Beijing Center
Center code:Beijing
Website:http://hgc.igtp.ac.cn
http://www.genomics.org.cn
Contact:hgc@igtp.ac.cn
----- Project Information
Center project name:1% project
Center clone name: RP11-202A21
----- Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator; ET 5% of reads
Chemistry: Dye-terminator Big Dye; 45% of reads
Assembly program: Phrap; version: 0.990329
Consensus quality: 832 bases at least Q40
Consensus quality: 1233 bases at least Q30
Consensus quality: 1299 bases at least Q20
Insert size: 1280; sum-of-contigs
Quality coverage: 1.10x in Q20 bases;sum-of-contigs

Location/Qualifiers
1. 171714
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"
/clone="RP11-202A21"
BASE COUNT 44183 a 38235 c 40748 g 48547 t 1 others
ORIGIN
Query Match 2.0%; Score 20; DB 88; Length 171714;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 813 gacagagctgcagcagacc 832
|||||
Db 136257 GACGAGCTGCAGCAGACC 136238
RESULT 17
AC017643 177816 bp DNA HTG 10-DEC-1999
LOCUS Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
DEFINITION pieces.
AC017643
AC017643.1 GI:6554354
VERSION
KEYWORDS
HTG: HTGS_PHASE2.
SOURCE
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 177816)
Adams,M. and Venter,J.C.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (10-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
This sequence was identified as CDM:10211585 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers
1. 177816
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
BASE COUNT 50527 a 38300 c 39085 g 49904 t
ORIGIN
Query Match 2.0%; Score 20; DB 65; Length 177816;

Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 50 ggcctcgtgctgacatgctg 69
|||||
Db 164707 GCGTCCTGCTGCATGTCG 164726
RESULT 18
AL512648 182365 bp DNA HTG 08-APR-2001
LOCUS Homo sapiens chromosome 13 clone RP11-465C24, *** SEQUENCING IN
DEFINITION PROGRESS ***, 7 unordered pieces.
AL512648
AL512648.6 GI:13568246
VERSION
KEYWORDS
HTG: HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 182365)
Submitted (08-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
On April 9, 2001 this sequence version replaced gi:13396721.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: ba465C24
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 98% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Consensus quality: 180306 bases at least Q40
Consensus quality: 181081 bases at least Q30
Consensus quality: 181448 bases at least Q20
Insert size: 181765; sum-of-contigs
Insert size: 165725; 8.2% error; agarose-fp
Quality coverage: 9.90x in Q20 bases; sum-of-contigs Quality
coverage: 11.02x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 10549: contig of 10549 bp in length
* 10550 10649: gap of 100 bp
* 10650 30196: contig of 19547 bp in length
* 30197 30296: gap of 100 bp
* 30297 32343: contig of 2047 bp in length
* 32344 32443: gap of 100 bp
* 32444 35679: contig of 3236 bp in length
* 35680 35779: gap of 100 bp
* 35780 53912: contig of 18133 bp in length
* 53913 54012: gap of 100 bp
* 54013 79432: contig of 25420 bp in length
* 79433 79532: gap of 100 bp
* 79533 182365: contig of 102833 bp in length.
Location/Qualifiers
1. 182365
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
FEATURES
source

```

misc_feature      /clone="RP11-465C24"
                  /clone_lib="RP11-11.2"
                  1..10549
misc_feature      /note="assembly_fragment:04706
                  fragment_chain:1"
                  10650..30196
misc_feature      /note="assembly_fragment:04716
                  fragment_chain:1"
                  30297..32343
misc_feature      /note="assembly_fragment:04843"
                  32444..35679
misc_feature      /note="assembly_fragment:04853"
                  35780..53912
misc_feature      /note="assembly_fragment:02763
                  fragment_chain:2"
                  54013..79432
misc_feature      /note="assembly_fragment:04938
                  fragment_chain:2"
                  79533..182365
misc_feature      /note="assembly_fragment:04959
                  fragment_chain:2
                  clone_end:77
                  vector_side:right"
BASE COUNT      53190 a 34137 c 35417 g 59008 t 613 others
ORIGIN

```

```

Query Match      2.0%; Score 20; DB 81; Length 182365;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 719 tcataccacacgctcgcg 738
|||||
Db 165448 TCATACCAACACGCTCTCG 165467

```

```

RESULT 19
AC007697 194897 bp DNA INV 10-MAR-2001
LOCUS Drosophila melanogaster, chromosome 2R, region 54D4-54E7, BAC clone
DEFINITION BACR12C23, complete sequence.
ACCESSION AC007697
VERSION AC007697.5 GI:13270541
KEYWORDS HTG.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 194897)

```

```

REFERENCE
AUTHORS Celiker,S.E., Adams,M.D., Krommiller,B., Tyler,D., Wan,K.H.,
Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
Rogers,J., An,H., Baldwin,D., Banton,J., Beeson,K.Y., Busam,D.A.,
Carlson,J.W., Center,A., Chemp,M., Davenport,L.B., Dietz,S.M.,
Dorison,K., Dorsett,V., Doup,L.E., Doyle,J.C., Dresner,D., Farfan,D.,
Ferdier,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
Gonzalez,M., Houch,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
Ibegam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
McInosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
Paciel,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
Stapleton,M., Strong,R., Svrlkas,R., Tector,C., Williams,S.M.,
Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
Sequencing of Drosophila chromosome 2R, region 54D4-54E7
TITLE Unpublished
JOURNAL 2 (bases 1 to 194897)
REFERENCE
AUTHORS Celiker,S.E., Agapayni,A., Arcaina,T.T., Baxter,E., Blazef,R.G.,
Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karia,K., Kearney,L.,
Klim,E., Lee,B., Lewis,S., Li,P., Lomocan,M.A., Mazda,P.,
Moshrefi,A.R., Moshrefi,M., Nixon,K., Paciel,J.M., Park,S.,
Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,

```

```

TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT On Mar 10, 2001 this sequence version replaced g1:5670564.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (http://www.fruitfly.org/sequence/) or send email
to bdg@fruitfly.berkeley.edu.
Location/Qualifiers:
1..194897
/organism="Drosophila melanogaster"
/strain="y: cn bw sp"
/db_xref="taxon:7227"
/chromosome="2R"
/map="54D4-54E7"
/clone="BACR12C23 (D665)"
/clone_lib="RP11-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcorI in
pBACE3.6)"

```

```

BASE COUNT      55759 a 41488 c 41712 g 55938 t
ORIGIN

```

```

Query Match      2.0%; Score 20; DB 4; Length 194897;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 50 ggcgtcgtgctgacatgctg 69
|||||
Db 128453 GCGTCCTGGCGACATGCTCG 128472

```

```

RESULT 20
AC022297 203195 bp DNA HTG 10-NOV-2000
LOCUS AC022297/c
DEFINITION Mus musculus clone RP23-336J1, WORKING DRAFT SEQUENCE, 28 unordered
pieces.
ACCESSION AC022297
VERSION AC022297.9 GI:11128306
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 203195)

```

```

REFERENCE
AUTHORS Metzger,M.L., Lewis,L.R., Hume,J., Edwards,C., Harris,C.,
Dederich,D., Thomas,S., Okwunonu,G., Carlock,C., Garner,T.,
Addison,S., Pace,A., Williams,G., Bonnin,D., Brooks,A., Brown,J.,
Hubay,C., Bunac,C., Burkett,C., Chacko,J., Chen,G., Chen,Z.,
Cox,C., Davis,C., Delgado,O., Ding,Y., Dugan-Rocha,S.,
Fernandez,C., Ferraguto,D., Fortum-Tansey,J., Gill,R.,
Gorell,J.H., Gunatirup,P., Haller,G., Hernandez,J., Hogues,M.,
Hosak,H., Hou,X., Huber,J., Jackson,L., Jia,Y., Kelly,J., Kelly,S.,
Kovar,C., Liu,J., Liu,W., Louesged,H., Lozado,R.J., Martin,R.,
Massey,E., McLeod,M.P., Mei,G., Moore,S., Morgan,M., Morris,S.,
Neal,D., Nelson,A., Nguyen,R., Nguyen,N., Ogun,M., Parish,B.,
Perez,L., Reiter,D., Say,J., Shen,H., Vasquez,L., Wallington,S.,
Williamson,A., Wrensford,G., Zhou,X., Bouck,J., Hodgson,A.,
Muzny,D.M., Rives,M., Scherer,S., Sodergren,E., Weinstein,G.,
Morley,K. and Gibbs,R.
Direct Submission
TITLE Unpublished
JOURNAL 2 (bases 1 to 203195)
REFERENCE
AUTHORS Morley,K.C.
TITLE Direct Submission

```

JOURNAL

COMMENT

Submitted (29-JAN-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 9, 2000 this sequence version replaced gi:11079345.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: MAF

Center clone name: RP23-336J1

Summary Statistics

Sequencing vector: M13; L08821

Chemistry: Dye-Primer Body: 84% of reads

Chemistry: Dye-terminator Big Dye: 16% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 178655 bases at least Q40

Consensus quality: 189131 bases at least Q30

Consensus quality: 193618 bases at least Q20

Estimated insert size: 196118; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-gel estimation

Quality coverage: 3.6x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently

consists of 28 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

```

1 38379: contig of 38379 bp in length
2 38479: gap of unknown length
3 38480: contig of 21207 bp in length
4 59686: gap of unknown length
5 59687: contig of 14515 bp in length
6 74301: gap of unknown length
7 74302: contig of 13799 bp in length
8 88200: gap of unknown length
9 88301: contig of 14506 bp in length
10 102806: gap of unknown length
11 102807: contig of 12666 bp in length
12 102907: gap of unknown length
13 115573: contig of 12666 bp in length
14 115573: gap of unknown length
15 127869: contig of 12197 bp in length
16 127870: gap of unknown length
17 127969: contig of 9603 bp in length
18 137572: gap of unknown length
19 137573: contig of 9885 bp in length
20 147558: gap of unknown length
21 147558: contig of 6459 bp in length
22 154116: gap of unknown length
23 154117: contig of 3959 bp in length
24 158175: gap of unknown length
25 158176: contig of 5675 bp in length
26 158275: gap of unknown length
27 163950: contig of 5010 bp in length
28 163951: gap of unknown length
29 164051: contig of 2248 bp in length
30 169060: gap of unknown length
31 169061: contig of 2050 bp in length
32 171409: gap of unknown length
33 171409: contig of 3283 bp in length
34 173559: gap of unknown length
35 173559: contig of 1410 bp in length
36 176942: gap of unknown length
37 176942: contig of 1804 bp in length
38 178452: gap of unknown length
39 178452: contig of 1804 bp in length
40 180356: gap of unknown length
41 180356: contig of 4076 bp in length
42 184532: gap of unknown length
43 184532: contig of 1754 bp in length
44 186383: gap of unknown length
45 186383: contig of 1754 bp in length

```

```

* 186386: gap of unknown length
* 186486: contig of 2714 bp in length
* 189200: gap of unknown length
* 189299: gap of unknown length
* 189300: contig of 2627 bp in length
* 191927: gap of unknown length
* 192026: gap of unknown length
* 192027: contig of 1418 bp in length
* 193445: gap of unknown length
* 193545: gap of 2167 bp in length
* 195712: contig of 2179 bp in length
* 195812: gap of unknown length
* 197991: contig of 2179 bp in length
* 198090: gap of unknown length
* 198091: gap of unknown length
* 200262: contig of 2171 bp in length
* 200362: gap of unknown length
* 201758: contig of 1397 bp in length
* 201759: gap of unknown length
* 201859: contig of 1337 bp in length.

```

FEATURES

source

location/qualifiers

1..203195

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="RP23-336J1"

BASE COUNT 55182 a 46344 c 46370 g 52543 t 2756 others

ORIGIN

Query Match 2.0%; Score 20; DB 67; Length 203195;

Best Local Similarity 100.0%; Pred. No. 8.9;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 803 aggcacagatgacagagctg 822

Db 69798 AGGCACAGATGACAGAGCTG 69779

RESULT 21

AC025210

LOCUS

DEFINITION

AC025210

AC025210

AC025210.2 GI:7321637

VERSION

HTGS: HTGS_PHASE1; HTGS_DRAFT.

KEYWORDS

SOURCE

ORGANISM

human.

Homo sapiens

Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 223589)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

JOURNAL
Submitted (07-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 24, 2000 this sequence version replaced gi:1188888.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Genome Center

Center code: MITR

Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center project name: L7839
Center clone name: 219_M.19

Summary Statistics

Sequencing vector: M13, M77815, 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 210657 bases at least Q40
Consensus quality: 217465 bases at least Q30
Consensus quality: 220046 bases at least Q20
Insert size: 213000; agarose-fp
Insert size: 221589; sum-of-ctrls
Quality coverage: 4.6 in Q20 bases; agarose-fp
Quality coverage: 4.4 in Q20 bases; sum-of-ctrls

* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

1 1249: contig of 1249 bp in length
1250 1349: gap of 100 bp
1350 4180: contig of 2831 bp in length
4181 4280: gap of 100 bp
4281 5454: contig of 1174 bp in length
5455 5554: gap of 100 bp
5555 7958: contig of 2404 bp in length
7959 8058: gap of 100 bp
8059 11369: contig of 3311 bp in length
11370 11469: gap of 100 bp
11470 13683: contig of 2214 bp in length
13684 13783: gap of 100 bp
13784 15932: contig of 2149 bp in length
15933 16032: gap of 100 bp
16033 19530: contig of 3498 bp in length
19531 19630: gap of 100 bp
19631 23177: contig of 3547 bp in length
23178 23277: gap of 100 bp
23278 28996: contig of 5719 bp in length
28997 29096: gap of 100 bp
29097 35377: contig of 6281 bp in length
35378 35477: gap of 100 bp
35478 43791: contig of 8314 bp in length
43792 43891: gap of 100 bp
43892 52630: contig of 8739 bp in length
52631 52730: gap of 100 bp
52731 68407: contig of 15677 bp in length
68408 68507: gap of 100 bp
68508 82738: contig of 14231 bp in length
82739 82838: gap of 100 bp
82839 98724: contig of 15886 bp in length
98725 98824: gap of 100 bp
98825 115023: contig of 16199 bp in length
115024 115123: gap of 100 bp
115124 134878: contig of 19755 bp in length
134879 134978: gap of 100 bp
134979 158924: contig of 23946 bp in length
158925 159024: gap of 100 bp
159025 191437: contig of 32413 bp in length

```

* 191438 191537: gap of 100 bp
* 191538 223589: contig of 32052 bp in length.
FEATURES
Source
1. .223589
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-219M19"
/clone_11b="RP11 Human Male BAC"
1. .1249
/note="assembly_fragment"
1350. .4180
/note="assembly_fragment"
4281. .5454
/note="assembly_fragment"
5555. .7958
/note="assembly_fragment"
8059. .11369
/note="assembly_fragment"
11470. .13683
/note="assembly_fragment"
13784. .15932
/note="assembly_fragment"
16033. .19530
/note="assembly_fragment"
19631. .23177
/note="assembly_fragment"
23278. .28996
/note="assembly_fragment"
29097. .35377
/note="assembly_fragment"
35478. .43791
/note="assembly_fragment"
43892. .52630
/note="assembly_fragment"
52731. .68407
/note="assembly_fragment"
68508. .82738
/note="assembly_fragment"
82839. .98724
/note="assembly_fragment"
98825. .115023
/note="assembly_fragment"
115124. .134878
/note="assembly_fragment"
134979. .158924
/note="assembly_fragment"
158925. .191437
/note="assembly_fragment"

BASE COUNT 61601 a 50789 c 49481 g 59714 t 2004 others
ORIGIN

Query Match 2.0%: Score 20; DB 69; Length 223589;
Best Local Similarity 100.0%: Pred. No. 8.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 813 gacagagctgacagagcc 832
|||||
Db 85662 GACAGAGCTGACAGAGCC 85681

RESULT 22
AE003802 262395 bp DNA INV 04-OCT-2000
LOCUS
DEFINITION Drosophila melanogaster genomic scaffold 142000013386047 section 41


```

gene
  /product="CT15697"
  /db_xref="FLYBASE:FBgn0004886"
  /db_xref="FLYBASE:FBgn0028382"
  <6206..>7308
  /gene="cyp33"
  /note="CG4886"
  /map="54C11-54D1"
  /db_xref="FLYBASE:FBgn0004886"
  /db_xref="FLYBASE:FBgn0028382"
  join(6213..6413,6478..6861,6918..7090,7164..7308)
  /gene="cyp33"
  /note="cyp33 gene product; Nucleotide sequence of the
  Celera sequence differs from the published sequence for
  this transcript"
  /codon_start=1
  /db_xref="FLYBASE:FBgn0004886"
  /db_xref="FLYBASE:FBgn0028382"
  /protein_id="AAF57839.1"
  /db_xref="GI:7302761"
  /translation="MSNDKRTIYVGLADEVTERLNNAPFIPGDIADIOMPADYESO
  RHGFATFEQSEDAIAIDNMNDELGRTRVNLAKPVKEDSPFIKADDMVL
  OHAGATLPBEGPEAEKVEYPTSTGPAVLEKAEKRNQVFFDRIKGNAGRIYMLLR
  AVVPTAEFRQLCTHEQGYGKGSFHRVLEPFQCGDFTNNNGTSGKSLYKFE
  NDENFLKHNSEGLTSMANGSANTNGSQCFFICTTKDMLDNKHVFGHVISGAEVVRK
  MERCGSKSTPSQKIVIVSCGELK"
  complement(join(7324..7910,8333..8405))
  /gene="Rpl18A"
  /note="Nucleotide sequence of the Celera sequence differs
  from the published sequence for this transcript."
  /product="CT20271"
  /db_xref="FLYBASE:FBgn0006510"
  /db_xref="FLYBASE:FBgn0010409"
  complement(<7324..>8405)
  /gene="Rpl18A"
  /note="CG6510"
  /map="54C11-54D1"
  /db_xref="FLYBASE:FBgn0006510"
  /db_xref="FLYBASE:FBgn0010409"
  complement(join(7395..7910,8333..8350))
  /gene="Rpl18A"
  /note="Rpl18A gene product; Nucleotide sequence of the
  Celera sequence differs from the published sequence for
  this transcript"
  /codon_start=1
  /db_xref="FLYBASE:FBgn0006510"
  /db_xref="FLYBASE:FBgn0010409"
  /protein_id="AAF57838.1"
  /db_xref="GI:7302760"
  /translation="MRAGKILKEVEYVGRKLPSKPEQTPLYKMRIFAPDNIYAKSRF
  WFLRLKKEKKTGTGIVSIKOYERSPKIKMFGILWLDSDSGTHNNYREYRLTV
  GCAVTCYRDMAHRARAHASIDIIKVDISIPAKTRRVHVKPHDSKIKFPLVQRVHH
  KGNRLKLSFRKPRFTTQ"

CDS
  Query Match      2.0%; Score 20; DB 5; Length 262395;
  Best Local Similarity 100.0%; Pred. No. 8.6;
  Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 ggcgtcgtgctgacatgctg 69
  |||||||||||||||||||
  Db 140906 GCGTCTGCTGACATGTCG 140925

RESULT 23
D89210/c  D89210      1152 bp      mRNA      PLN      13-MAR-1998
LOCUS      Schizosaccharomyces pombe mRNA, partial cds, clone: SY 1164.
ACCESSION  D89210
VERSION      D89210.1  GI:1749627
KEYWORDS
SOURCE      Schizosaccharomyces pombe (strain:PR745) cDNA to mRNA,
ORGANISM      clone_lib:library of H. Nojima clone:SY1164.
              Schizosaccharomyces pombe
              Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

```

```

Schizosaccharomycetales; Schizosaccharomycetaceae;
Schizosaccharomycetes.
REFERENCE
  1 (bases 1 to 1152)
  yoshioaka,S.
  TITLE      Direct Submission
  JOURNAL
  Submitted (15-NOV-1996) to the DDBJ/EMBL/GenBank databases. Sachio
  Yoshioaka, Tsukita Cell Axis Project ERATO JST, Kyoto Research Park;
  17 Chudouji Minamimachi, Shimokyo-ku, Kyoto, Kyoto 600, Japan
  (E-mail: yoshio@cell.tsukita.jst.go.jp, Tel: +81-75-315-7913,
  fax: +81-75-315-6420)
  2 (sites)
  Yoshioaka,S., Kato,K., Nakai,K., Okayama,H. and Nojima,H.
  Identification of open reading frames in Schizosaccharomyces pombe
  cDNAs
  JOURNAL
  DNA Res. 4 (6), 363-369 (1997)
  MEDLINE
  96162722
  FEATURES
    source
      1..1152
      /organism="Schizosaccharomyces pombe"
      /strain="PR745"
      /db_xref="taxon:4896"
      /clone="SY1164"
      /clone_lib="library of H. Nojima"
      <1..522
      /note="similar to Saccharomyces cerevisiae ORF YGL136C,
      EMBL Accession Number Z72658"
      /codon_start=1
      /protein_id="BAA13871.1"
      /db_xref="GI:1749628"
      /translation="INKEDVYIDVGFAPGSMSQVAKKLVGNKGVIGIDIQHTAPPEG
      VLEPIQDIRDPTLTFLFALRLHEPNTSDICRVDAVAVIDMLHKATGIRIPIHA
      LSEILCASALHVALVFLKSGSPICKFYMDDEDADIDNLLKSHFRVQVKKPRASLKE
      SRKAYVCLERK"

BASE COUNT      325 a      211 c      257 g      359 t
ORIGIN

Query Match      1.9%; Score 19; DB 14; Length 1152;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 ctgcataccaccctgcttc 251
  |||||||||||||||||||
  Db 853 CTGCATCACCCTGCTTTC 835

RESULT 24
AF273691      1279 bp      mRNA      ROD      22-FEB-2001
LOCUS      Mus musculus ribosome receptor isoform mRrp2 mRNA, complete cds,
DEFINITION  alternatively spliced.
ACCESSION  AF273691
VERSION      AF273691.1  GI:13094692
KEYWORDS
SOURCE      house mouse.
ORGANISM      Mus musculus.
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
              Kim,Y.-J., Lee,M.-C., Kim,S. and Chun,J.-Y.
              Identification and characterization of multiple isoforms of a mouse
              ribosome receptor
              Gene 261 (2), 337-344 (2000)
              JOURNAL
              PUBMED
              11167022
              REFERENCE
              2 (bases 1 to 1279)
              Kim,Y.-J., Lee,M.-C., Kim,S. and Chun,J.-Y.
              Direct Submission
              JOURNAL
              Submitted (01-JUN-2000) Kumbo life and Environmental Science
              Laboratory, 572 Sangam-Dong, Kwangsan-Gu, Kwangju 506-712, Korea
              1..1279
              Location/Qualifiers
              /organism="Mus musculus"
              /strain="ICR"

```

```

CDS
    /db_xref="taxon:10090"
    /dev_stage="13.5 day embryo"
    210..962
    /note="alternatively spliced"
    /codon_start=1
    /product="ribosome receptor isoform mRrp2"
    /protein_id="AAK11971.1"
    /db_xref="GI:13094693"
    /translation="MDIYDTQTLGVVFGFVNVSAIGIFLVSTFSMKETSYEALAN
    ORKEMAKTHHOKGKKKKETVEKKKKEKKKEKNGKIPEDHLDPNVTIILKEPVV
    SAVAAPTSVSSVGHPTLATYVAMPQEKIASPPKDRKKKEKVAKVPASVYNSI
    OYLAASKSAILENTPEKVPVNAVPPVSGKASSPATSSQAKKVDAAANOGKSSMAPAQC
    QKASNVQSEAPKODAPAKKSGSKKGEPC"
    795..854
    /note="Region: tandem repeat"
    misc_feature
    /note="Region: tandem repeat"
    357 a 309 c 346 g 267 t
    BASE COUNT
    ORIGIN

Query Match
Best Local Similarity 1.9%; Score 19; DB 94; Length 1279;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 115 cagcagccagcgaacaaa 133
    |||||||
Db 791 CAGCAGCCAGCCCAAAAA 809

RESULT 25
AE005021/C 10128 bp DNA BCT 12-FEB-2001
LOCUS Halobacterium sp. NRC-1 section 52 of 170 of the complete genome.
DEFINITION
ACCESSION AE005021 AE004437
VERSION AE005021.1 GI:10580357
KEYWORDS
SOURCE
    Halobacterium sp. NRC-1.
    Halobacterium sp. NRC-1.
    Archaea: Euryarchaeota: Halobacteriales: Halobacteriaceae:
    Halobacterium.
REFERENCE
    1 (bases 1 to 10128)
    Ng,M.V., Kennedy,S.P., Mahairas,G.G., Berquist,B., Pan,M.,
    Shukla,H.D., Lasky,S.R., Balliga,N., Thorsson,V., Sproga,J.,
    Swartzell,S., Weir,D., Hall,J., Dahl,T.A., Welti,R., Goo,Y.A.,
    Leitnauer,B., Keller,K., Cruz,R., Danson,M.J., Hough,D.W.,
    Maddocks,D.G., Jablonski,P.E., Krebs,M.P., Angevine,C.M., Dale,H.,
    Isenbarger,T.A., Peck,R.F., Pohlischod,M., Spudich,J.L.,
    Jung,K.-H., Alam,M., Freitas,T., Hou,S., Daniels,C.J., Dennis,P.P.,
    Omer,A.D., Ehardt,H., Lowe,T.M., Liang,P., Riley,M., Hood,L. and
    DasSarma,S.
    From the cover: genome sequence of halobacterium species NRC-1
    Proc. Natl. Acad. Sci. USA 97 (22), 12176-12181 (2000)
11016950
2 (bases 1 to 10128)
    Ng,M.V., Kennedy,S.P., Mahairas,G.G., Berquist,B., Pan,M.,
    Shukla,H.D., Lasky,S.R., Balliga,N., Thorsson,V., Sproga,J.,
    Swartzell,S., Weir,D., Hall,J., Dahl,T.A., Welti,R., Goo,Y.A.,
    Leitnauer,B., Keller,K., Cruz,R., Danson,M.J., Hough,D.W.,
    Maddocks,D.G., Jablonski,P.E., Krebs,M.P., Angevine,C.M., Dale,H.,
    Isenbarger,T.A., Peck,R.F., Pohlischod,M., Spudich,J.L.,
    Jung,K.-H., Alam,M., Freitas,T., Hou,S., Daniels,C.J., Dennis,P.P.,
    Omer,A.D., Ehardt,H., Lowe,T.M., Liang,P., Riley,M., Hood,L. and
    DasSarma,S.
    Direct Submission
    Submitted (14-JUL-2000) Institute for Systems Biology, 4225
    Roosevelt Way NE, Seattle, WA 98105, USA
FEATURES
    source
    1..10128
    /organism="Halobacterium sp. NRC-1"
    /strain="NRC-1"
    /db_xref="taxon:64091"
    /complement(818..2119)
    /gene="VNG0782H"
    /complement(818..2119)
    CDS

```

```

    /gene="VNG0782H"
    /note="hypothetical protein"
    /codon_start=1
    /transl_table=11
    /product="Vng0782h"
    /protein_id="AAG19250.1"
    /db_xref="GI:10580358"
    /translation="MSDGEDEPAETTEREAGEAADTTAVDPDTLESRLDDEALEAAS
    TESDLDDEASLDAVAADIEAELPEPDDEESREALTJDLAEVRGVDQKGPYA
    SDAETVAASSTIGETRWTOGDEDDLAAYVAQPEEAALDGETFDETGDLGDLA
    VSEITSAIGLADADADPTIESTVAATDSLDGCVADAOEMDLSVNRKLRAGCYEVMG
    DKHDFPEPTALAKMEORGVNEYKILLCEHLDSEEFQRICLDLHVAKPALDVA
    GQLARDPKALEVVGKIGVEDGEIEEYVAEDSPMLRTTALKAVETIGSRSTOV
    VADDLVAERADVRQAARALGIGTTRAIAPIVDLTLEDSEETVRAASAAALVOIGTE
    AALEAAEYADRAVYVEQEAATKATVATLDRHDAADASTGSA"
    complement(2172..2849)
    /gene="psa"
    /note="VNG0784C"
    complement(2172..2849)
    /gene="psa"
    /note="psa"
    /codon_start=1
    /transl_table=11
    /product="CDP-diacylglycerol-serine
    O-phosphatidyltransferase"
    /protein_id="AAG19251.1"
    /db_xref="GI:10580359"
    /translation="MOPREVGRLGVAADVTTANALGFLAVVTAAPSNRRLAIVLA
    AIDGLDGVVARRLGSGTPAGKRLDLSLAVASRGVAPACIIVAATGNTMOQAPGIGA
    LVNGGPAASAVRGLITATDTGHEHREGVQTTLAATLLAAGVLAGVEQTYVIGA
    GVLAVLVSTIEYPDLLVRLDALMGVAVGAVAPALAGAPKALFEMAGATLVFAP
    RPYWRAA"
    2946..3395
    /gene="pcy"
    /note="VNG0786C"
    2946..3395
    /gene="pcy"
    /note="pcy"
    /codon_start=1
    /transl_table=11
    /product="plastocyanin homology"
    /protein_id="AAG19252.1"
    /db_xref="GI:10580360"
    /translation="MHRRAFLAGGTLSVGLACIGPSVSKSDYDGMVNGEVPQE
    PVEGEDMATHTQIADGTIVTANTGSRNHTVAYAGSLDEGASYPASGGFDEPARRA
    MWDNEGGSNIPGATVEVTLTPVEYYFCIPHBGMRKLVYDD"
    complement(3456..4076)
    /gene="rps3e"
    /note="VNG0787C"
    complement(3456..4076)
    /gene="rps3e"
    /note="rps3e"
    /codon_start=1
    /transl_table=11
    /product="30S ribosomal protein S3e"
    /protein_id="AAG19253.1"
    /db_xref="GI:10580361"
    /translation="MSERSKSGKODEKRMVTVLAPEEPFRAELGETLAEEPDOYDRT
    IOVALTDDGSDNNIKILFPOIDVGSASQFOALTRDYOGLVRSKVAAT
    VYATLTDYRVRIOPVATYTKOAOOSQOHAIRTMIDIVEAGEERTFALLNSIIEG
    RLSSAIYDEANTITPLRVEYEKATLEHPEVHEEETAVDFSGE"
    complement(4073..4234)
    /gene="VNG0788H"
    complement(4073..4234)
    /gene="VNG0788H"
    /note="hypothetical protein"
    /codon_start=1
    /transl_table=11
    /product="Vng0788h"
    /protein_id="AAG19254.1"
    /db_xref="GI:10580362"
    /translation="MTTVDDTGTGVTVTRIERADTAGMSTVDYVVALTVAAVTAOT
    ANQHTTQOS"

```

```
gene complement(4315. .5418)
      /gene="VNG0789C"
      /complement(4315. .5418)
CDS /gene="VNG0789C"
      /note="conserved hypothetical protein"
      /codon_start=1
      /transl_table=11
      /product="Vng0789c"
      /protein_id="AAG19255.1"
      /db_xref="GI:10580363"
      /translation="MATIGRSNDSDGVRDVAGVSGAEFVRYTAAPGCGVVAAGLL
AGVCDGVSPQVRRVSEPPSHSGADGSOAAPTASHVTCGEAAVAASVGVADA
FALAREVGTGHPDPVTAGAVADVGPAPDARAAGAWTRRPGVPTADLAAGLAS
TIVHASFSGDEQAGAFIAELGLPEPPDAAARRRLASVALDRTADAPPCAAADVDA
LRPRATPDGMMATVEGADVLYNVAAMAAGHGVALGLVADRTAVLAAMAHAPPSHA
AVRAASLARHRGVYADVPTASTVARLLRDYATPVALAVSTETALATTTDADAAH
VLEPGSVTRRAVTTLDSSDADAVIEIVRGGL"
gene complement(5423. .5890)
      /gene="rps15p"
      /note="VNG0790G"
      /complement(5423. .5890)
CDS /gene="rps15p"
      /note="rps15p"
      /codon_start=1
      /transl_table=11
      /product="30S ribosomal protein S15P"
      /protein_id="AAG19256.1"
      /db_xref="GI:10580364"
      /translation="MARHTRRRGSSGSDRPAPADEPPESDVDEDAIEERVELAEOG
HPSQIGLKRDEGVGTVPDVKATGKVEILPAHDAEPLPEDFRLLEKAVRL
HEHEVMGODHNRKALQNTQSKIRRLADYRGDKIDEFPVSYETAREIIEE"
      6115. .6186
      /gene="trn11"
      /product="trnA-Ala"
      6115. .6186
      /gene="trn11"
      /complement(6477. .8846)
gene /gene="trt6"
      /note="VNG0793G"
      /complement(6477. .8846)
CDS /gene="trt6"
      /note="trt6"
      /codon_start=1
      /transl_table=11
      /product="Htr6 transducer"
      /protein_id="Htr6_19257.1"
      /db_xref="GI:10580365"
      /translation="MSDGIITGALRNAASVAVAPDAIRRGAKFAVFIYVLYIAGAG
VFAFOSTETVEHOTTRQLAESNLEADAGAMEGOQTHARSVSODEALRODRRAP
YILKQOLPADVSMHLYNETRGTVASTELAIEGRSLAEIDA PMTTAAVPGPGND
SAVMATEQSYRSPVNLDERVMAFASVPRREGAHLVYTRIQVQPRLSANSTRPT
IINTGDEPVLDSNSRFDAGALDSTIAVRNDTTATTVTSNGSVYLSATPHWDWTY
TSVDTGCAFAVRDTGQTVYLVLAIVLAISLAIGIALGRHTVPLRLARRADIISG
TFDVIDLSTRIDIEVGRLYGSFDMRVSIDRLIOEAAVAEANAARAEELRTDIEDG
AQAEEARAKATAEASERLOERADYSSEVMAVADDLTERLEDDEDEEMRAVEEF
NAMLDLEATIAOVAGEFADVDETLOVATGAEITTSQTVSERIQELIADGALIOOH
DIERAGEMDELASATQVAAASATVAETRAADAVEGEGRDAESRIDMATIESIS
ADAVIDIILQERMSDIGITIEITIDIACTNMLANANTEARAKRDGGFVAVNE
VVDLAETKQAAADIESEIQAVQAEETVADIRATSEHIDGVSTVERAAAIEDTV
DAIEDANSHGIOEISDATEDOQATQSVVRVDVADISQHTVEDAQCVSAAAEQCSAS
VAIEARSADDLRBDALATVYVQFETRADADEPDADTTVDASADDTGD"
      complement(8843. .9611)
gene /gene="yufn"
      /note="VNG0794G"
      /complement(8843. .9611)
CDS /gene="yufn"
      /note="yufn"
      /codon_start=1
      /transl_table=11
      /product="ABC transporter (lipoprotein)"
      /protein_id="AAG19258.1"
      /db_xref="GI:10580366"
      /translation="WYVALGGLDRSFNDANRGAQRARLDHGIAVTYHNEPDSIAELG
```

```
ETQATLAASNPAVDLVCVGFROADGLESAANAAPDOOFMIVDTAVDASNVAAYEE
EOTGSPQAGRALALITTRRIDAGGATNPARTVGVGVDEPLTFGFPQAGVYTH
TASDIDVLAAYTGFDEVDENARAAANMAYAGADITYHHAAGGAGVYLQAAQNGRYA
IGVSDOSRNSNPYADVNLASMKRITDVAIDYDIEATIDBDLADGVVSLGLATNGVS
VQGTQDGPAPVAVDVALGETREALVAGDITVTPERPQASAGGGV"
BASE COUNT 1353 a 3628 c 3425 g 1722 t
ORIGIN

Query Match 1.9%; Score 19; DB 1; Length 10128;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 386 cgcgcgcacacgacgagc 404
|||||
Db 574 CGCGGCCACGACGAGGTC 556

RESULT 26
AR083116/c 11873 bp DNA PAT 01-SEP-2000
LOCUS AR083116
DEFINITION Sequence 32 from patent US 5976803.
ACCESSION AR083116
VERSION AR083116.1 GI:10009906
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 11873)
AUTHORS Meek, K.D.
TITLE Genetic test for equine severe combined immunodeficiency disease
JOURNAL Patent: US 5976803-A 32 02-NOV-1999;
FEATURES
SOURCE 1. 11873
/organism="unknown"
BASE COUNT 3585 a 2320 c 2669 g 3299 t
ORIGIN

Query Match 1.9%; Score 19; DB 9; Length 11873;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 actctgcgcgtcgtcat 229
|||||
Db 6446 ACTTCTGCGCTGCTGCA 6428

RESULT 27
AR083115/c 11878 bp DNA PAT 01-SEP-2000
LOCUS AR083115
DEFINITION Sequence 31 from patent US 5976803.
ACCESSION AR083115
VERSION AR083115.1 GI:10009905
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 11878)
AUTHORS Meek, K.D.
TITLE Genetic test for equine severe combined immunodeficiency disease
JOURNAL Patent: US 5976803-A 31 02-NOV-1999;
FEATURES
SOURCE 1. 11878
/organism="unknown"
BASE COUNT 3586 a 2323 c 2668 g 3301 t
ORIGIN

Query Match 1.9%; Score 19; DB 9; Length 11878;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 211 actctgcgcgtcgtcat 229
 Db 6446 ACTTCTGCCGCTGTCAT 6428

RESULT 28
 AR083114/c 11883 bp DNA PAT 01-SEP-2000
 LOCUS AR083114
 DEFINITION Sequence 28 from patent US 5976803.
 ACCESSION AR083114
 VERSION AR083114.1 GI:10009904
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 11883)
 AUTHORS Meek, K.D.
 TITLE Genetic test for equine severe combined immunodeficiency disease
 JOURNAL Patent: US 5976803-A 28 02-NOV-1999;
 FEATURES Location/Qualifiers
 source 1..11883
 /organism="unknown"
 BASE COUNT 3588 a 2324 c 2669 g 3302 t
 ORIGIN

Query Match 1.9%; Score 19; DB 9; Length 11883;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 actctgcgcgtcgtcat 229
 Db 6449 ACTTCTGCCGCTGTCAT 6431

RESULT 29
 SPBC262 36493 bp DNA PLN 23-MAY-1999
 LOCUS S. pombe chromosome II cosmid c262.
 DEFINITION AL022103
 ACCESSION AL022103.1 GI:2956767
 VERSION
 KEYWORDS 60s ribosomal protein l16-c; beta transducin; beta-adaptin; cell division protein l16; homolog; deoxycytidylate deaminase; glycylypeptide n-tetradecanoyltransferase; histidyl-trna synthetase; l12lp; mannose-6-phosphate isomerase; phosphatidylinositol phosphate phosphatase; popl; protein transport protein; rpl16c; sun family protein; tetrahydrofolate synthase; transmembrane transporter; WD repeat protein; WD-40 repeat.

SOURCE Schizosaccharomyces pombe
 ORGANISM Schizosaccharomycetes; Schizosaccharomycetes; Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes; Schizosaccharomycetales; Schizosaccharomycetaceae; Schizosaccharomycetes.

REFERENCE 1 (bases 1 to 36493)
 AUTHORS Wood, V., Rajandream, M.A., Barrell, B.G., Devlin, K. and Churcher, C.M.
 TITLE Direct Submission
 JOURNAL Submitting (04-MAR-1998) European Schizosaccharomycetes genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk

COMMENT Details of yeast sequencing at the Sanger Centre are available on the World Wide Web.
 (URL, <http://www.sanger.ac.uk/projects/S.pombe/>)
 During 1995 to 1996 about 80% of S. pombe chromosome 1 was sequenced by the Sanger Centre. The sequencing of the S. pombe genome is now being continued with funding from The European Commission. Fourteen European sequencing laboratories, including the Sanger Centre, are participating in the project.
 Protein coding regions (CDS) have been predicted with the help of computer analysis using the Genefinder program in Pombase (an ACEDB database) with additional predictions for the branch-acceptor sites supplied by the program Sp3splice. CAUTION: It is possible that for any individual CDS we may have underestimated or overestimated the

number of introns/exons or we may not have chosen the correct splice donor/acceptor sites.
 CDS are numbered using the following system eg SPBC25H2.01c. SP (S. pombe), B (chromosome 2), c25H2 (cosmid name), .01 (first CDS), c (complementary strand).
 The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous.
 The length in codons is given for each CDS.
 IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because the entire sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions.
 Cosmid c262 overlaps 5' with c484.

FEATURES

source

1..36493
 /organism="Schizosaccharomyces pombe"

/strain="972h-"

/db_xref="taxon:4896"

/chromosome="II"

/clone="cosmid c262"

/map="IIR"

complement(1..1156)

/gene="SPBC262.01c"

/note="1121"

1..1709

/note="9nominal overlap with SPBC484 S. pombe chromosome 2"

complement(1..1156)

/partial

/gene="SPBC262.01c"

/function="transmembrane transporter; affects cell cycle progression"

/note="SPBC262.01c, len:>383"

/codon_start=1

/label=1121

/product="transmembrane transporter l12lp."

/protein_id="CA17881.1"

/db_xref="GI:2956768"

/db_xref="SPTREMBL:O43000"

/translation="MALNLRKTFSPYGLNKVEOKLIDKIDPILISCCVSEFINVLRSSINNAVLSGMEDELMKGNLEODINVTGCIITIGOLGSAVALORVAPRAMEVNMILGLMTITSPFAVSHYRALMILREFPAVAEASTFACHTITLAWKYESLCKRAGIETSAAGLVETPMAGIQTAVHSSLNKGGSLSMKRWFTIDGLITLPLSLGLEPDDVETTPAPVETPEDEKELSPKRLPARPKKPLTLAKIDIVSRWRTGLTWLFEETQALAVNLMQOMKMKNSKFSVAQINNPVTAVGVSTLGASVSDKLGNPWFEGFLLCVITVTATITLAWNPDCAKFPFAFASCTYAGQAWFMSANDICRND"

2337..5567

/gene="SPBC262.02"

2337..5567

/note="SPBC262.02, len:1075, SIMILARITY: Saccharomyces cerevisiae, Q12271, YOR109y, phosphatidylinositol phosphate phosphatase (1107 aa), fasta scores: opt: 1630, E(-): 0, (39.0% identity in 1100 aa)"

/codon_start=1

/label=SPBC262.02

/product="phosphatidylinositol phosphate phosphatase; synaptojanin-like"

/protein_id="CA17882.1"

/db_xref="GI:2956769"

/db_xref="SPTREMBL:O43001"

/translation="WQCLLRKPNRLAVLNKDHLMFHSVPQNKNSLVCVAEFTALS EKLPEGRKISHRVYGLGLESGNLFCLVIGSASVAVRERKERVIRMEVCFYS VRSNMWDIRROENSPDIDPDYDTGDSYKVAEPFSLRKLITNGSFYSLDID ITTRRLQTSOTMTEPOVDSMTQPMNMEFRLRIKFRSHLNGDEKSGALDGRFTC AIRGPASREQKRLGQTRLSLISLSTSLAGCTPRLSGDDGVAVFERTLIDS SKVCSTQVQKSTITFEWEGVQWFGKIDITTSLETRAPAEKHTFLSLEEGPVH IINLGTSGERSLSERLRQHIQSPEDLIHLTEFEDHNSOIRSEDENKRTPMITYSD AETFGVFENNEGOSIVQDVFRTNCDICDRVNIQNLVSRYFLSEQVMYTRINAG YDEMVOHSTIVANNGDALARIYTGGALESFTRKGLSIGALNLSVGRIMYN FODRGROETIDLLGLRIDQHPVILYDPIHLYVNHLEKRNRESEKHNVAIPASV LNGCSATTKLENLFPETPLADIIYVGFQETVQVTPQVYISADPARKREMEVCYRL LNKCTSGPGIVQLRSGLVGTALVIFKESCLPSIKNVEGTAKTKGLGVSNGKAV

```

AIFRDEYTGCLPITSHLAGYNTYNDERDHYRTTASGLNFRGRSIFNHDYVWVEGD
FNRIJLTYEEVVPICIAQGLKSYLFEYDOLNKMQLKVPFSELPTEPPYKFKDI
GTDIYDTSKDHBPAMTDRIYLGELVPSYGVSPYSHRPIVATYENAIYKVBRE
KKKILFEEILYNOKKOEVRDASQSTYLIDAGSVACKPNILPILPANGDKIKPSSE
RKMWFDDGICPAKSIAPAPPEPTRLNPSRPIINFEPTAEEDWISNTKOSFDRKSSLD
STPALSPPBSILARSSVSORSSTSIILPKPNPTPDHLVAVRVPFLPPEPRGSSSS
GVPAPMUTPVNPPPTPPKSSASORSGLDASPRESSISMKPLV"
complement(join(5906, .6005, 6056, .6187, 6252, .6313,
6364, .6378))
/gene="SPBC262.03c"
complement(join(5906, .6005, 6056, .6187, 6252, .6313,
6364, .6378))
/gene="SPBC262.03c"
/note="SPBC262.03c, len:101, SIMILARITY: Saccharomyces
cerevisiae, S6B2_YEAST, protein transport protein sec61
beta 2 subunit, (88 aa), fasta scores: opt: 242,
E():2.1e-09, (46.38 identity in 82 aa)"
/codon_start=1
/label="SPBC262.03c"
/product="putative protein transport protein; yeast sec61
beta 2 subunit-like"
/protein_id="CAI17883.1"
/db_xref="GI:2956770"
/translation="MSSTKASGSVKNSAASAPGPKQIRRAAVEKNTESNSGPAG
ARAAGAPGSTPLIKLYTDEAGCFKVDPVVWVLSGVFASVFLHIVARILKKFAPSE
"
misc_feature
complement(6006, .6020)
/note="ctaacgtatctag, splice branch and acceptor"
complement(6050, .6055)
/note="gtgaagc, splice donor sequence"
complement(6188, .6210)
/note="ctaacctttgtgtgattgtag, splice branch and
acceptor"
complement(6246, .6251)
/note="gttcgt, splice donor sequence"
complement(6314, .6324)
/note="ctaacatagtag, splice branch and acceptor"
complement(6358, .6363)
/note="gtgaact, splice donor sequence"
complement(6875, .7463)
/gene="SPBC262.04c"
complement(6975, .7463)
/gene="SPBC262.04c"
/note="SPBC262.04c, len:161, SIMILARITY: Saccharomyces
cerevisiae, YIF1_YEAST, required for maintenance of
mitochondrial DNA, (145 aa), fasta scores: opt: 357,
E():7.5e-17, (44.08 identity in 141 aa)
sequence possibly needs truncating to second met"
/codon_start=1
/label="SPBC262.04c"
/product="hypothetical protein"
/protein_id="CAI17884.1"
/db_xref="GI:2956771"
/translation="MLRALGSRLLVASRPAAYRFSOOLRPVRFIHKMSKTPINS
PKRSSGPNQAIKNGAVIYSGOIPVANGKVEGTGDTROCLINLIDGVLDEAGSS
LKTIYKVNIFLADMDFPAVANKYITVLEDPKPARSCVAVKTYPLSTGCKIETCIA
LE"
gene
join(8051, .8091, 8150, .8702)
/gene="SPBC262.05"
/note="rp116c"
join(8051, .8091, 8150, .8702)
/gene="SPBC262.05"
/note="SPBC262.05, len:196"
/codon_start=1
/label="rp116c"
/product="60s ribosomal protein 116-c."
/protein_id="CAI17885.1"
/db_xref="GI:2956772"
/translation="MSFQKLVIIIDAKGHLMGLASTVAKOLLAGQVVVRCETINI
SCHFFRNKLKYLAVLRKACRYNPSRGAFHFRAPSRIETFAVRGMPLHKTTFGNIALKN

```

```

LQALEGIPPEFDKRLVVPALRVLRILKPSRKYCTIGRLSEVGMKYNKIYKLEER
RKIKSAFYQAKANQKHIVAKTKSSVNEKLAVFG"
8092, .8097
/note="gtatga, splice donor sequence"
8135, .8149
/note="ctaacgtatctag, splice branch and acceptor"
8361, .8432
/gene="SPBC262.05"
/note="P500783 Ribosomal protein L13 signature"
complement(join(8859, .10245, 10250, .10408, 10453, .10698,
10752, .10797, 10835, .11043, 11090, .11116))
/gene="SPBC262.06c"
complement(join(8859, .10245, 10290, .10408, 10453, .10698,
10752, .10797, 10835, .11043, 11090, .11116))
/gene="SPBC262.06c"
/note="SPBC262.06c, len:676, SIMILARITY: Drosophila
melanogaster, Q24253, beta-adaptin drosophila 1., (921
aa), fasta scores: opt: 1724, E():0, (45.28 identity in
aa)
Query Match 1.9% Score 19; DB 15; Length 36493;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 233 ctgcatacaccctgttc 251
|||||
Db 27668 CTCGATCACCCCGCTTTC 27686
RESULT 30
AC015691 67442 bp DNA HTG 17-NOV-1999
AC015691 LOCUS
DEFINITION Homo sapiens chromosome 11 clone RP11-54F1 map 11, LOW-PASS
SEQUENCE SAMPLING.
AC015691 AC015691.1 GI:6446996
VERSION HTG; HTGS_PHASE0; NULL.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 67442)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 11, clone RP11-54F1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 67442)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Batra,N., Beckerly,R., Boguslavsky,L., Bouknight,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hages,B., Heathford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., MacDonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Strange-Rhoman,N., Stojanovic,N., Subramanian,A., Telamas,J.,
Testaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
Project Information
Center project name: L1351
Center clone name: 54_F_1

```

* NOTE: This record contains 78 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely

* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

| | | | | | |
|---|-----------------------------------|---|-------|-----------------------------------|---|
| 1 | 867: contig of 867 bp in length | * | 28407 | 29224: contig of 818 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 868: contig of 805 bp in length | * | 29225 | 30081: contig of 857 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 1673: contig of 841 bp in length | * | 30082 | 30974: contig of 893 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 2514: contig of 884 bp in length | * | 30975 | 31915: contig of 941 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 3398: contig of 896 bp in length | * | 31916 | 32758: contig of 843 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 4294: contig of 870 bp in length | * | 32759 | 33645: contig of 887 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 5164: contig of 860 bp in length | * | 33646 | 34475: contig of 830 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 6023: contig of 911 bp in length | * | 34476 | 35388: contig of 913 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 6024: contig of 911 bp in length | * | 35389 | 36231: contig of 843 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 6935: contig of 881 bp in length | * | 36232 | 37036: contig of 805 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 7816: contig of 873 bp in length | * | 37037 | 37908: contig of 872 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 8689: contig of 899 bp in length | * | 37909 | 38748: contig of 840 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 9588: contig of 842 bp in length | * | 38749 | 39648: contig of 900 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 10430: contig of 871 bp in length | * | 39649 | 40515: contig of 867 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 11301: contig of 871 bp in length | * | 40516 | 41344: contig of 829 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 12172: contig of 876 bp in length | * | 41345 | 42174: contig of 830 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 13048: contig of 863 bp in length | * | 42175 | 43005: contig of 831 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 13911: contig of 856 bp in length | * | 43006 | 43846: contig of 841 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 14767: contig of 861 bp in length | * | 43847 | 44739: contig of 893 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 15628: contig of 903 bp in length | * | 44740 | 45640: contig of 901 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 16531: contig of 779 bp in length | * | 45641 | 46543: contig of 903 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 17310: contig of 855 bp in length | * | 46544 | 47389: contig of 846 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 18165: contig of 901 bp in length | * | 47390 | 48276: contig of 887 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 19066: contig of 840 bp in length | * | 48277 | 49160: contig of 884 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 19906: contig of 854 bp in length | * | 49161 | 50039: contig of 879 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 20760: contig of 792 bp in length | * | 50040 | 50865: contig of 826 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 21552: contig of 828 bp in length | * | 50866 | 51712: contig of 847 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 22380: contig of 856 bp in length | * | 51713 | 52571: contig of 859 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 23236: contig of 905 bp in length | * | 52572 | 53408: contig of 837 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 24141: contig of 867 bp in length | * | 53409 | 54227: contig of 819 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 25008: contig of 813 bp in length | * | 54228 | 55090: contig of 863 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 25821: contig of 845 bp in length | * | 55091 | 55971: contig of 881 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 26666: contig of 925 bp in length | * | 55972 | 56863: contig of 892 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 27591: contig of 816 bp in length | * | 56864 | 57747: contig of 884 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | 28406: contig of 816 bp in length | * | 57748 | 58623: contig of 876 bp in length | * |
| * | gap of unknown length | * | * | gap of unknown length | * |
| * | | * | 58624 | 59480: contig of 857 bp in length | * |
| * | | * | * | gap of unknown length | * |
| * | | * | 59481 | 60365: contig of 885 bp in length | * |
| * | | * | * | gap of unknown length | * |

* gap of unknown length
* 60366 61266: contig of 901 bp in length
* gap of unknown length
* 61267 62134: contig of 868 bp in length
* gap of unknown length
* 62135 62991: contig of 857 bp in length
* gap of unknown length
* 62992 63853: contig of 862 bp in length
* gap of unknown length
* 63854 64832: contig of 979 bp in length
* gap of unknown length

Query Match 1.9%: Score 19; DB 63; Length 67442;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 767 gaagagacatgaacaact 785
|||||
Db 54512 GAAGAGACATGAACAACACT 54530

RESULT 31
AC011428 79970 bp DNA HTG 15-NOV-2000
LOCUS Homo sapiens chromosome 5 clone Xp1-380H11, WORKING DRAFT
DEFINITION AC011428
SEQUENCE, 4 unordered pieces.
AC011428.3 GI:11178060
VERSION HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 79970)
DOE Joint Genome Institute.
Unpublished
2 (bases 1 to 79970)
DOE Joint Genome Institute.
Direct Submission
Submitted (06-OCT-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Nov 15, 2000 this sequence version replaced gi:6604400.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 1189149, H122
Center clone name: Xp1-380H11

Summary Statistics
Consensus quality: 76910 bases at least Q40
Consensus quality: 78432 bases at least Q30
Consensus quality: 78894 bases at least Q20
Estimated insert size: 72000; pulse field gel estimation
Quality coverage: 7.19 in Q20 bases; pulse field gel estimation
Quality coverage: 6.49 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 3224: contig of 3224 bp in length
* 3225 3324: gap of unknown length
* 3325 13887: contig of 10563 bp in length
* 13888 13987: gap of unknown length
* 13988 32432: contig of 18445 bp in length

* 32433 32532: gap of unknown length
* 32533 79970: contig of 47438 bp in length.
FEATURES
Source
1. 79970
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="XXp1-380H11"

BASE COUNT 20846 a 18934 c 18691 g 21162 t 337 others
ORIGIN

Query Match 1.9%: Score 19; DB 62; Length 79970;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 776 atgaacaactctgaacct 794
|||||
Db 50545 ATGAACAACCTCTGACACT 50563

RESULT 32
AC068903 82033 bp DNA ROD 20-SEP-2000
LOCUS Mus musculus chromosome 2 clone ct7-196d19 strain 129/Sv ES cell
DEFINITION AC068903
line CJ7, complete sequence.
AC068903.11 GI:9837949
VERSION HTG.
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 82033)
Shauli, S., Rahnal, R., Yao, Z., Rithidech, K. and Roe, B.A.
Mus musculus Chromosome 2 BAC Clone Cltb-196d19
Unpublished
2 (bases 1 to 82033)
Shauli, S., Rahnal, R., Yao, Z., Rithidech, K. and Roe, B.A.
Direct Submission
Submitted (11-MAY-2000) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 82033)
Shauli, S., Rahnal, R., Yao, Z., Rithidech, K. and Roe, B.A.
Direct Submission
Submitted (20-SEP-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Aug 17, 2000 this sequence version replaced gi:9797790.

Project Information
Center Project Name: 1189149, H122
Center clone name: Xp1-380H11

Summary Statistics
Consensus quality: 76910 bases at least Q40
Consensus quality: 78432 bases at least Q30
Consensus quality: 78894 bases at least Q20
Estimated insert size: 72000; pulse field gel estimation
Quality coverage: 7.19 in Q20 bases; pulse field gel estimation
Quality coverage: 6.49 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

BASE COUNT 23822 a 16503 c 15652 g 26056 t
ORIGIN
1. 82033
/organism="Mus musculus"
/strain="129/Sv ES cell line CJ7"
/db_xref="taxon:10090"
/chromosome="2"
/clone="ct7-196d19"
/clone_lib="Caltech CTRB-BAC library"

Query Match 1.9%: Score 19; DB 94; Length 82033;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaacacaaggaagaagaa 555
|||||
Db 61880 GAACACAAGGAGAAGAA 61862

RESULT 33
AC068952/c

LOCUS AC068952 95292 bp DNA 18-AUG-2000
DEFINITION Mus musculus chromosome 2 clone ct7-254f1 strain 129/Sv ES cell
line C37, complete sequence.
ACCESSION AC068952
VERSION AC068952.10 GI:9845112
KEYWORDS HTG.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 95292)
AUTHORS Shull, S., Rahhal, R., Yao, Z., Ritchiech, K. and ROE, B.A.
TITLE Mus musculus Chromosome 2 BAC Clone c17b-254f1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 95292)
AUTHORS Shull, S., Rahhal, R., Yao, Z., Ritchiech, K. and ROE, B.A.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2000) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 3 (bases 1 to 95292)
AUTHORS Shull, S., Rahhal, R., Yao, Z., Ritchiech, K. and ROE, B.A.
TITLE Direct Submission
JOURNAL Submitted (18-AUG-2000) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Aug 18, 2000 this sequence version replaced gi:9795634.
FEATURES
source
1..95292
Location/Qualifiers
/organism="Mus musculus"
/strain="129/Sv ES cell line C37"
/db_xref="taxon:10090"
/chromosome="2"
/clone="ct7-254f1"
/clone_lib="Caltech CTB-BAC library"
BASE COUNT 28084 a 18825 c 18104 g 30279 t
ORIGIN
Query Match 1.9%; Score 19; DB 94; Length 95292;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 537 gaaccacacagagaaga 555
Db 76659 GAACCAACAGGAGAAGA 76641.
RESULT 34
AL354832_1
WPCOMMENT
Sequence split into 4 fragments LOCUS AL354832 Accession AL354832
Fragment Name Begin End
AL354832_0 1 110000
AL354832_1 100001 210000
AL354832_2 200001 310000
AL354832_3 300001 372955
Continuation (2 of 4) of AL354832 from base 100001 (AL354832 Homo sapiens chromosome 20
Query Match 1.9%; Score 19; DB 79; Length 110000;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 215 ctgcgcgtcgtcaccc 233
Db 66912 CTGCCGCTGCTGATCCC 66930
RESULT 35
AC024591 111071 bp DNA HTG 18-JUL-2000
LOCUS AC024591/c
DEFINITION Homo sapiens chromosome 16 clone RP11-511G21, WORKING DRAFT

SEQUENCE, 13 ordered pieces.
AC024591
AC024591.3 GI:9256450
VERSION
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 111071)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 111071)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (29-FEB-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 18, 2000 this sequence version replaced gi:7705016.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 602754
Center clone name: RPCI-11_511G21

Summary Statistics
Consensus quality: 104021 bases at least Q40
Consensus quality: 108348 bases at least Q30
Consensus quality: 109319 bases at least Q20
Estimated insert size: 118930; agarose-fp estimation
Estimated insert size: 110521; sum-of-ctnigs estimation
Quality coverage: 8.44 in Q20 bases; agarose-fp estimation
Quality coverage: 9.08 in Q20 bases; sum-of-ctnigs estimation.
* NOTE: This is a "working draft" sequence. It currently
* consists of 13 ctnigs. Gaps between the ctnigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1
4000 3999: contig of 3999 bp in length
4099: gap of unknown length
4100 6224: contig of 2125 bp in length
6225 6325: gap of unknown length
6325 14602: contig of 8278 bp in length
14603 14703: gap of unknown length
14703 54245: contig of 39543 bp in length
54246 54346: gap of unknown length
54346 58569: contig of 4224 bp in length
58570 58670: gap of unknown length
58670 59480: contig of 811 bp in length
59481 59580: gap of unknown length
59580 60683: contig of 1103 bp in length
60683 60784: gap of unknown length
60784 62950: contig of 2167 bp in length
62951 63050: gap of unknown length
63050 64092: contig of 1042 bp in length
64093 64193: gap of unknown length
64193 85067: contig of 20875 bp in length
85068 85167: gap of unknown length
85168 91149: contig of 5982 bp in length
91150 91250: gap of unknown length
91250 108236: contig of 1687 bp in length
108237 108336: gap of unknown length
108337 111071: contig of 2735 bp in length.
Location/Qualifiers
1..111071
/organism="Homo sapiens"
/db_xref="taxon:9606"

| BASE COUNT | ORIGIN |
|------------|---------------------------------------|
| 27517 a | /chromosome="16" |
| 27152 c | /clone="Rp11-511G21" |
| 28256 g | /clone.lib="RC1 human BAC library 11" |
| 26944 t | 1202 others |

| | | | | |
|--------------------------|--------|--------------|----------|---------------|
| Query Match | 1.9% | Score 19 | DB 68 | Length 111071 |
| Best Local Similarity | 100.0% | Pred. No. 35 | | |
| Matches 19, Conservative | 0 | Mismatches 0 | Indels 0 | Gaps 0 |

| | | | |
|----|-------|-------------------|-------|
| OY | 280 | gccgcgcgcatcgcgtg | 298 |
| | | | |
| Db | 19460 | gccggcgccgatcgctg | 19442 |

| | | |
|------------|---|--|
| RESULT | 36 | |
| AC020962 | | |
| LOCUS | AC020962 | |
| DEFINITION | Mus musculus clone RP21-126H5, WORKING DRAFT SEQUENCE, 21 unordered pieces. | |

| | |
|-----------|------------------------------|
| ACCESSION | AC020962 |
| VERSION | AC020962.1 |
| KEYWORDS | HTG; HTGS_PHASE1; HTGS_DRAFT |
| SOURCE | house mouse. |
| ORGANISM | Mus musculus |

REFERENCE
1 (bases 1 to 128562)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus

AUTHORS DOE Joint Genome Institute
TITLE Sequencing of Mouse
JOURNAL unpublished

| | |
|-----------|------------------------------|
| REFERENCE | 2 (bases 1 to 128562) |
| AUTHORS | DOE Joint Genome Institute |
| TITLE | Direct Submission |
| JOURNAL | Submitted (12-JAN-2000) Proc |

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center

Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

```

Consensus quality: 113514 bases at least Q40
Consensus quality: 131055 bases at least Q30
Consensus quality: 136285 bases at least Q20
Estimated insert size: 128565: sum-of-contigs estimation
Estimated insert size: 122000: pulse field gel estimation
Quality coverage: 7.89x in Q20 bases; pulse field gel estimation
Quality coverage: 7.48x in Q20 bases; sum-of-contigs estimation
=====

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

| | | | |
|---|------|--------|-----------------------------|
| * | 1 | 1198: | contig of 1198 bp in length |
| * | | | gap of unknown length |
| * | 1199 | 2531: | contig of 1333 bp in length |
| * | | | gap of unknown length |
| * | 2532 | 4140: | contig of 1609 bp in length |
| * | | | gap of unknown length |
| * | 4141 | 5328: | contig of 1188 bp in length |
| * | | | gap of unknown length |
| * | 5329 | 6673: | contig of 1345 bp in length |
| * | | | gap of unknown length |
| * | 6674 | 8037: | contig of 1364 bp in length |
| * | | | gap of unknown length |
| * | 8038 | 9583: | contig of 1546 bp in length |
| * | | | gap of unknown length |
| * | 9584 | 11046: | contig of 1453 bp in length |

| BASE COUNT | ORIGIN | FEATURES | SOURCE |
|------------|--------|------------------------------|------------|
| * | | gap of unknown length | |
| * | 11047 | contig of 1795 bp in length | |
| * | 12842 | gap of unknown length | |
| * | 14057 | contig of 1215 bp in length | |
| * | 15422 | gap of unknown length | |
| * | 16974 | contig of 1365 bp in length | |
| * | 18747 | gap of unknown length | |
| * | 20177 | contig of 1552 bp in length | |
| * | 21644 | gap of unknown length | |
| * | 22956 | contig of 1773 bp in length | |
| * | 24396 | gap of unknown length | |
| * | 26494 | contig of 1430 bp in length | |
| * | 46983 | gap of unknown length | |
| * | 78190 | contig of 1467 bp in length | |
| * | 128562 | gap of unknown length | |
| * | 128562 | contig of 1312 bp in length | |
| * | 128562 | gap of unknown length | |
| * | 128562 | contig of 1440 bp in length | |
| * | 128562 | gap of unknown length | |
| * | 128562 | contig of 2098 bp in length | |
| * | 128562 | gap of unknown length | |
| * | 128562 | contig of 20489 bp in length | |
| * | 128562 | gap of unknown length | |
| * | 128562 | contig of 31207 bp in length | |
| * | 128562 | gap of unknown length | |
| * | 128562 | contig of 50373 bp in length | |
| | | Location:Qualifiers | |
| | | 1..128562 | |
| | | /organism="Mus musculus" | |
| | | /db_xref="taxon:10090" | |
| | | /clone="RP21-126H5" | |
| 31926 a | 3261 c | 31576 g | 32269 t |
| | | | 130 others |

| | | | | |
|----------------------|-----------------|---------------|-----------|----------------|
| Query Match | 1.9% | Score 19; | DB 66; | Length 128562; |
| Best Local Similarly | 100.0%; | Pred. No. 34; | | |
| Matches 19; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |

```
QY      280 gccgcgagcgatcgcgtg 298  
          |||||  
Db 117029 GCCGCGCGCATCGGCTG 117047
```

| RESULT | 37 |
|------------|---|
| AC073882 | |
| LOCUS | AC073882 |
| DEFINITION | AC073882 130734 bp DNA |
| DEFINITION | Mus musculus chromosome 2 clone ct7-305f12 strain 129/SvEv Cell |
| DEFINITION | line Cj7, complete sequence. |
| ACCESSION | U0073882 |

| | |
|-----------|--------------------|
| ACCESSION | AC073882 |
| VERSION | AC073882.4 |
| KEYWORDS | HTG. |
| SOURCE | house mouse. |
| ORGANISM | mouse Mus musculus |

| ORGANISM | MUS MUSCULUS |
|---|--------------|
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus | |
| 1 (bases 1 to 10721) | |

| REFERENCE | AUTHORS | TITLE |
|-----------------------|---|--|
| 1 (bases 1 to 150734) | Jia, H., Lin, S., Rithideck, K. and Roe, B.A. | Mus musculus Chromosome 2 BAC Clone ct7-305f12 |

| JOURNAL REFERENCE | AUTHORS | TITLE |
|-------------------|---|-------------------|
| unpublished | 2 (bases 1 to 130734) | |
| | Jia, H., Lin, S., Rithideck, K. and Roe, B.A. | Direct Submission |

Submitted (01-JUL-2000) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, U.S.A

| REFERENCE AUTHORS TITLE | 3 (bases 1 to 130734) Jia, H., Lin, S., Ritchie, K. and Roe, B. A. Direct Submission |
|-------------------------------|--|
| | |

JOURNAL
Submitted (04-MUG-2000) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
4 (bases 1 to 130734)

AUTHOR: Jia, H., Lin, S., Rithideck, K. and Roe, B. A.
 TITLE: Direct Submission
 JOURNAL: Submitted (24-NOV-2000) Department of Chemistry and Biochemistry,
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 COMMENT: On Aug 4, 2000 this sequence version replaced gi:9280733.
 FEATURES: Location/Qualifiers
 source 1. 130734
 /organism="Mus musculus"
 /strain="129/SVES Cell Line C17"
 /db_xref="taxon:10090"
 /chromosome="2"
 /clone="ctf-305f12"
 /clone_lib="Caltech CTFB-BAC library"
 BASE COUNT 41025 a 25364 c 26535 g 37652 t
 ORIGIN
 Query Match 1.9%; Score 19; DB 94; Length 130724;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 537 gaaccacacaggaagaaga 555
 ||||||||||||||||
 Db 65070 GAACCAACAGGAGAGAA 65088
 RESULT 38
 HSJ1057D4 131888 bp DNA PRI 01-MAR-2001
 LOCUS HSJ1057D4
 DEFINITION Human DNA sequence from clone RP5-1057D4 on chromosome 20 contains
 a spermidine synthase (SPSY) pseudogene, a CpG island, ESTs, STS
 and GSSs, complete sequence.
 AL121777
 AL121777.39 GI:10862842
 HMG: CPG island; spermidine synthase.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 131888)
 Blakey, S.
 Direct Submission
 Submitted (26-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humgeny@sanger.ac.uk
 requests: clonerequests@sanger.ac.uk
 On Oct 17, 2000 this sequence version replaced gi:10716550.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 validation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 EMBL, SW, SWISSPROT, Tr., TREMBL, Wp., WORMEP, Information
 on the WORMEP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormep
 This sequence was generated from part of bacterial clone contigs of human
 chromosome 20, constructed by the Sanger Centre Chromosome 20
 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/RGP/Chr20
 This sequence is the entire insert of clone RP5-1057D4 The true
 left end of clone RP1-73E16 is at 71240 in this sequence. The true
 right end of clone CTD-2653D5 is at 84358 in this sequence. This
 sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; the
 assembly was confirmed by restriction digest. RP5-1057D4 is from
 the library RPCT-5 constructed by the group of Pieter de Jong. For

```

further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pcYpac2.
FEATURES
Source
Location/Qualifiers
1. .131888
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RPS-1057D4"
/clone_11b="RcT-5"
313. .930
/note="L2 repeat: matches 2043. .2749 of consensus"
1411. .1543
/note="MIR repeat: matches 80. .222 of consensus"
1623. .1934
/note="AluY repeat: matches 1. .311 of consensus"
2162. .2215
/note="L2 repeat: matches 2696. .2749 of consensus"
2444. .2646
/note="L2 repeat: matches 2353. .2521 of consensus"
2929. .3407
/note="match: GSS: Em:A0520865"
3104. .3214
/note="L2 repeat: matches 2629. .2746 of consensus"
3638. .3839
/note="MIR repeat: matches 32. .256 of consensus"
3916. .4051
/note="AluX repeat: matches 1: .123 of consensus"
4052. .4341
/note="AluY repeat: matches 1. .295 of consensus"
4342. .4477
/note="AluX repeat: matches 123. .298 of consensus"
5436. .5600
/note="AluX repeat: matches 170. .299 of consensus"
5601. .5892
/note="AluY repeat: matches 1. .291 of consensus"
5893. .6050
/note="AluX repeat: matches 1. .170 of consensus"
6179. .6230
/note="MIR repeat: matches 206. .256 of consensus"
6186. .6565
/note="L2 repeat: matches 2292. .2731 of consensus"
6567. .6606
/note="L2 repeat: matches 2148. .2189 of consensus"
6607. .6913
/note="AluX repeat: matches 1. .307 of consensus"
6914. .7005
/note="L2 repeat: matches 2064. .2148 of consensus"
7104. .7559
/note="MIR repeat: matches 31. .512 of consensus"
8768. .8969
/note="MER20 repeat: matches 1. .216 of consensus"
9696. .10011
/note="L2 repeat: matches 2357. .2734 of consensus"
11637. .12065
/note="MIR repeat: matches 20. .531 of consensus"
12302. .12377
/note="MIR repeat: matches 115. .190 of consensus"
13520. .13783
/note="L2 repeat: matches 2432. .2749 of consensus"
13953. .14155
/note="MIR repeat: matches 267. .468 of consensus"
14420. .14705
/note="AluY repeat: matches 1. .294 of consensus"
complement(14545. .15063)
/note="match: GSS: Em:A0713490"
complement(14589. .15060)
/note="match: GSS: Em:A0760291"
complement(14602. .15072)
/note="match: GSS: Em:A085065"
15070. .15576
/note="match: GSS: Em:A0709181"
15088. .15403

```

```

/misc_feature      /note="match: GSS: Em:AQ198342"
15088. .15583
/note="match: GSS: Em:AQ680811"
/misc_feature      15094. .15492
/note="match: GSS: Em:AQ082327"
repeat_region      15423. .15656
/note="L2 repeat: matches 1609. .1854 of consensus"
repeat_region      15716. .15894
/note="MER20 repeat: matches 4. .187 of consensus"
repeat_region      15898. .16155
/note="L2 repeat: matches 1915. .2159 of consensus"
repeat_region      16156. .16457
/note="Alusx repeat: matches 3. .304 of consensus"
repeat_region      16458. .16568
/note="L2 repeat: matches 2159. .2263 of consensus"
repeat_region      16569. .16865
/note="Alusx repeat: matches 1. .301 of consensus"
repeat_region      16870. .17097
/note="L2 repeat: matches 2263. .2531 of consensus"
repeat_region      17103. .17406
/note="Alusx repeat: matches 1. .301 of consensus"
repeat_region      18176. .18380
/note="MIR repeat: matches 36. .260 of consensus"
repeat_region      19512. .19745
/note="MIR repeat: matches 13. .261 of consensus"
/misc_feature      /note="match: GSS: Em:AQ383848"
21138. .23316
/note="CPG island"
/misc_feature      /evidence="not_experimental"
21792. .22251
/note="230 copies 2 mer cc 55% conserved"
misc_feature        21814. .22282
/note="single clone region. Assembly confirmed by
restriction digest."
repeat_region      21829. .22122
/note="6 copies 49 mer 64% conserved"
repeat_region      21832. .22251
/note="12 copies 35 mer 57% conserved"
repeat_region      21934. .22123
/note="5 copies 38 mer 67% conserved"
repeat_region      22000. .22257
/note="6 copies 43 mer 62% conserved"
repeat_region      23418. .23568
/note="MIR repeat: matches 18. .187 of consensus"
repeat_region      23981. .24040
/note="MER33 repeat: matches 233. .293 of consensus"
repeat_region      24041. .24328
/note="Alub repeat: matches 1. .288 of consensus"
repeat_region      24329. .24543
/note="MER33 repeat: matches 23. .233 of consensus"
repeat_region      25195. .25242
/note="24 copies 2 mer tg 72% conserved"
repeat_region      27239. .27425
/note="MIR repeat: matches 55. .249 of consensus"
repeat_region      27638. .27889
/note="LMC4 repeat: matches 7694. .7977 of consensus"
repeat_region      27900. .28343
/note="L2 repeat: matches 2292. .2750 of consensus"
repeat_region      28434. .28597
/note="MIR repeat: matches 57. .232 of consensus"
repeat_region      29051. .29345
/note="Aluio repeat: matches 2. .297 of consensus"
repeat_region      29456. .29717
/note="Alusx repeat: matches 1. .254 of consensus"
repeat_region      29873. .30045
/note="MLTIC repeat: matches 6. .163 of consensus"
repeat_region      30046. .30348
/note="Alusx repeat: matches 1. .306 of consensus"
repeat_region      30349. .30598
/note="MLTIC repeat: matches 163. .464 of consensus"
repeat_region      30721. .30798
/note="MER5B repeat: matches 44. .117 of consensus"

```

```

repeat_region      30802. .30927
/note="LMB8 repeat: matches 6040. .6168 of consensus"
repeat_region      30942. .31001
/note="MER5B repeat: matches 113. .173 of consensus"
repeat_region      31509. .31809
/note="Alusx repeat: matches 1. .304 of consensus"
repeat_region      32898. .33202
/note="Alusg repeat: matches 1. .310 of consensus"
repeat_region      33556. .33701
/note="MER5A repeat: matches 1. .146 of consensus"

Query Match      1.9%; Score 19; DB 93; Length 131888;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      215 ctgcgcgtcgtcatcccc 233
Db      22024 CTGCCGCTGCTGCATCCCC 22042

RESULT 39
AC005738/c      134506 bp      DNA      PRI      20-OCT-1998
LOCUS      Homo sapiens chromosome 5, BAC clone 7g12 (LBNL H126), complete
DEFINITION
ACCESSION      AC005738 AC004619 LB1405 LB1406 LB1678 LB1679 LB1680 LB1860 LB1861
LB1862 AC001042 AC001043 AC001520
VERSION      AC005738.1 GI:3687213
KEYWORDS      HTG.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 134506)
AUTHORS      Kimmery,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
Kadner,K., Miguel,T., Miller,C., Piltuck,S., Pollard,M.,
Rojeski,H., Subramanian,S. and Martin,C.H.
TITLE      Submitted (01-OCT-1998) Human Genome Center, DOE Joint Genome
Institute, Lawrence Berkeley National Laboratory, MS 74-157,
Berkeley, CA 94720, U.S.A.
JOURNAL      4 (bases 1 to 134506)
Ricke,D.O.
REFERENCE      Direct Submission
AUTHORS      Submitted (13-OCT-1998) Los Alamos National Laboratory, DOE Joint
Genome Institute, TA43, HRL-1, LS-3, MS M888, Los Alamos, NM 87545,
USA
JOURNAL      5 (bases 1 to 134506)
Ricke,D.O.
REFERENCE      Direct Submission
AUTHORS      Submitted (20-OCT-1998) Los Alamos National Laboratory, DOE Joint
Genome Institute, TA43, HRL-1, LS-3, MS M888, Los Alamos, NM 87545,
USA
JOURNAL      Sequence submitted by:
COMMENT      DOE Joint Genome Institute.
FEATURES      Location/Qualifiers
source      1. .134506
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="5p"
/chromosome="5"
join(1. .29755,78169. .82395,123646. .134506)
source

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5p"
/clone="BAC 7612"
/note="LBNT_H126"
complement(748..891)
/note="GRAIL 2 excellent exon, frame 2"
misc_feature
2635..2694
/note="GRAIL 2 excellent exon, frame 2"
repeat_region
complement(2883..2994)
/rpt_family="MLT1"
complement(3940..4205)
repeat_region
/rpt_family="Alu"
complement(5036..5148)
/rpt_family="MIR"
9194..9650
/rpt_family="MERA2"
9745..10340
/rpt_family="MERA2"
10346..10624
/rpt_family="Alu"
10674..10876
/rpt_family="MERA2"
12173..12472
/rpt_family="Alu"
14779..14955
/rpt_family="L1"
15163..15487
/rpt_family="Alu"
complement(19591..19738)
/rpt_family="MIR"
20251..20549
/rpt_family="Alu"
complement(21622..21774)
/note="GRAIL 2 excellent exon, frame 1"
misc_feature
complement(22628..22797)
/rpt_family="MIR"
complement(23374..23671)
/rpt_family="Alu"
24507..24641
/rpt_family="MIR"
28893..60995
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5p"
/clone="PI 210H6"
/note="LBNT_H52"
complement(34480..34672)
/rpt_family="MLT1"
complement(34676..34821)
/note="GRAIL 2 excellent exon, frame 0"
repeat_region
complement(34837..35181)
/rpt_family="THE1"
38746..38854
/note="GRAIL 2 excellent exon, frame 1"
repeat_region
39274..39561
/rpt_family="Alu"
39856..40341
/rpt_family="MIR"
complement(40375..40451)
/rpt_family="MLT1"
44067..44382
/rpt_family="Alu"
44592..44900
/rpt_family="Alu"
45213..45338
/rpt_family="L1"
45376..45708
/rpt_family="Alu"
46954..47235
/rpt_family="Alu"

repeat_region
50715..51019
/rpt_family="Alu"
51522..51600
/rpt_family="MIR"
complement(51960..52100)
/rpt_family="MIR"
complement(52323..52389)
/rpt_family="MIR"
complement(52403..52692)
/rpt_family="Alu"
complement(52752..53008)
/rpt_family="Alu"
complement(53050..53388)
/rpt_family="Alu"
complement(53346..53568)
/rpt_family="MIR"
complement(55100..55354)
/rpt_family="L1"
complement(55478..55533)
/rpt_family="Alu"
complement(56376..56704)
/rpt_family="L1"
57025..57175
/rpt_family="MLT1"
complement(58128..58416)
/rpt_family="Alu"
complement(58603..58884)
/rpt_family="L1"
59416..59747
/rpt_family="MIR"
join(60081..78733,82162..124917)
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PI 380H11"
/chromosome="5"
/map="5p"
/note="LBNT_H122"
complement(60553..60710)
/rpt_family="Alu"
complement(60718..60914)
/rpt_family="THE1"
complement(61908..62249)
/rpt_family="THE1"
complement(62279..63781)
/rpt_family="MSTAR"
complement(63786..64160)
/rpt_family="THE1"
65491..65778
/rpt_family="Alu"
complement(66977..67113)
/note="GRAIL 2 excellent exon, frame 2"
67119..67157
/rpt_family="THE1"
complement(68013..68120)
/note="GRAIL 2 excellent exon, frame 1"
68546..68651
/rpt_family="MIR"
68894..69226
/rpt_family="THE1"
complement(69684..70023)
/rpt_family="THE1"
70567..70881
/rpt_family="Alu"
complement(72274..72537)
/rpt_family="Alu"
complement(74449..74802)
/note="GRAIL 2 excellent exon, frame 2"
complement(77047..77436)
/rpt_family="MLT1"
79647..79829
/note="GRAIL 2 excellent exon, frame 2"
complement(82876..82937)
/note="GRAIL 2 excellent exon, frame 0"

```

```
repeat_region      83530. .83674
                    /rpt_family="Alu"
repeat_region      complement(87731. .87819)

Query Match        1.9%; Score 19; DB 86; Length 134506;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      776 atgacacatctgaact 794
        |||
Db      88571 ATGACACATCTGAACT 88553

RESULT  40
AC023898/c
LOCUS   AC023898.6 141198 bp  DNA  HTG  04-NOV-2000
DEFINITION Mus musculus clone RP23-281H12, *** SEQUENCING IN PROGRESS ***, 68
AC023898
VERSION  AC023898.6 GI:11079363
KEYWORDS HTG; HTGS; PHASEL.
SOURCE    house mouse.
ORGANISM  Mus musculus
REFERENCE
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
          1 (bases 1 to 141198)
          Metzger M.L., Lewis L.R., Hume J., Edwards C., Harris C.,
          Dedrich D., Thomas S., Okwuonu G., Carlock C., Garner T.,
          Addison S., Pace A., Williams G., Bonnin D., Brooks A., Brown J.,
          Bahay C., Bunac C., Burkett C., Chacko J., Chen G., Chen Z.,
          Cox C., Davis C., Delgado O., Ding Y., Dugan-Rocha S.,
          Fernandez C., Ferraguto D., Forcum-Tansey J., Gill R.,
          Gorrell J.H., Gunaratne P., Haller G., Hernandez J., Hoques M.,
          Hosak H., Hou X., Huber J., Jackson L., Jia Y., Kelly J., Kelly S.,
          Kovar C., Liu D., Liu W., Louised H., Lozado R.J., Martin R.,
          Massey E., McLeod M.P., Mel G., Moore S., Morgan M., Morris S.,
          Neal D., Nelson A., Nguyen R., Nguyen N., Ogub M., Parish B.,
          Perez L., Reiter D., Say J., Shen H., Vasquez L., Wallington S.,
          Williamson A., Wrensford G., Zhou X., Bouck J., Hodgson A.,
          Muzny D.M., Rives M., Scherer S., Sodergren E., Weinstock G.,
          Worley K. and Gibbs R.
          Direct Submission
          Unpublished
          2 (bases 1 to 141198)
          Worley K.C.
          Direct Submission
          Submitted (20-FEB-2000) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA
          On Nov 3, 2000 this sequence version replaced gi:8248594.
          ----- Genome Center
          Center: Baylor College of Medicine
          Center code: BCM
          Web site: http://www.hgsc.bcm.tmc.edu/
          Contact: hgsc-help@bcm.tmc.edu
          ----- Project Information
          Center project name: MABD
          Center clone name: RP23-281H12
          ----- Summary Statistics
          Sequencing vector: M13; L08821
          Chemistry: Dye-terminator Big Dye; 10% of reads
          Assembly program: Phrap; version 0.990329
          Consensus quality: 96352 bases at least Q40
          Consensus quality: 119814 bases at least Q30
          Consensus quality: 130654 bases at least Q20
          Estimated insert size: 106211; sum-of-contigs estimation
          Quality coverage: 0x in 020 bases; agarose-gel estimation
          Quality coverage: 1.5x in 020 bases; sum-of-contigs estimation
          -----
          * NOTE: Estimated insert size may differ from sequence length
          * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
          * NOTE: This is a 'working draft' sequence. It currently
```

```
* consists of 68 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1      2727      2726: contig of 2726 bp in length
*      2827      2826: gap of unknown length
*      6808      6807: contig of 3981 bp in length
*      6908      6907: gap of unknown length
*      10328     10327: contig of 3420 bp in length
*      10428     10427: gap of unknown length
*      14552     14551: contig of 4124 bp in length
*      14652     14651: gap of unknown length
*      17639     17638: contig of 2987 bp in length
*      17739     17738: gap of unknown length
*      23093     23092: contig of 5354 bp in length
*      23193     23192: gap of unknown length
*      25971     25970: contig of 2778 bp in length
*      26071     26070: gap of unknown length
*      28816     28815: contig of 2745 bp in length
*      28916     28915: gap of unknown length
*      32009     32008: contig of 3093 bp in length
*      32109     32108: gap of unknown length
*      35880     35879: contig of 3771 bp in length
*      35980     35979: gap of unknown length
*      39176     39175: contig of 3196 bp in length
*      39276     39275: gap of unknown length
*      41990     41989: contig of 2714 bp in length
*      42090     42089: contig of unknown length
*      44450     44449: contig of 2360 bp in length
*      44550     44549: gap of unknown length
*      46328     46327: contig of 1778 bp in length
*      46428     46427: gap of unknown length
*      48944     48943: contig of 2516 bp in length
*      49044     49043: gap of unknown length
*      51172     51171: contig of 2128 bp in length
*      51272     51271: gap of unknown length
*      53834     53833: contig of 2562 bp in length
*      53934     53933: gap of unknown length
*      56124     56123: contig of 2190 bp in length
*      56224     56223: gap of unknown length
*      58222     58221: contig of 1998 bp in length
*      58322     58321: gap of unknown length
*      60196     60195: contig of 1874 bp in length
*      60296     60295: gap of unknown length
*      63109     63108: contig of 2813 bp in length
*      63209     63208: gap of unknown length
*      64983     64982: contig of 1774 bp in length
*      65083     65082: gap of unknown length
*      68031     68030: contig of 2948 bp in length
*      68131     68130: gap of unknown length
*      69914     69913: contig of 1783 bp in length
*      70014     70013: gap of unknown length
*      72171     72170: contig of 2157 bp in length
*      72271     72270: gap of unknown length
*      73627     73627: gap of 1357 bp in length
*      73728     73727: gap of unknown length
*      76148     76147: contig of 2420 bp in length
*      76248     76247: gap of unknown length
*      78973     78972: contig of 2725 bp in length
*      79073     79072: gap of unknown length
*      81459     81458: contig of 2386 bp in length
*      81559     81558: gap of unknown length
*      83164     83163: contig of 1605 bp in length
*      83264     83263: gap of unknown length
*      84726     84725: contig of 1462 bp in length
*      84826     84825: gap of unknown length
*      86931     86931: contig of 2106 bp in length
*      87032     87031: gap of unknown length
*      89044     89044: contig of 2013 bp in length
*      89144     89144: gap of unknown length
```

89145 90654: contig of 1510 bp in length
 * 90655 90754: gap of unknown length
 * 90755 93175: contig of 2421 bp in length
 * 93176 93275: gap of unknown length
 * 93276 95024: contig of 1749 bp in length
 * 95025 95124: gap of unknown length
 * 95125 96172: contig of 1048 bp in length
 * 96173 96272: gap of unknown length
 * 96273 97373: contig of 1465 bp in length
 * 97378 97837: gap of unknown length
 * 97838 99532: contig of 1695 bp in length
 * 99533 99632: gap of unknown length
 * 99633 100907: contig of 1275 bp in length
 * 100908 101007: gap of unknown length
 * 101008 102521: contig of 1514 bp in length
 * 102522 102621: gap of unknown length
 * 102622 103668: contig of 1047 bp in length
 * 103669 103768: gap of unknown length
 * 103769 104929: contig of 1161 bp in length
 * 104930 105029: gap of unknown length
 * 105030 106714: contig of 1685 bp in length
 * 106715 106814: gap of unknown length
 * 106815 108598: contig of 1784 bp in length
 * 108599 108698: gap of unknown length
 * 108699 109927: contig of 1229 bp in length
 * 109928 110027: gap of unknown length
 * 110028 111479: contig of 1452 bp in length
 * 111480 111579: gap of unknown length
 * 111580 112613: contig of 1034 bp in length
 * 112614 112713: gap of unknown length
 * 112714 113746: contig of 1033 bp in length
 * 113747 113846: gap of unknown length
 * 113847 115329: contig of 1483 bp in length
 * 115330 115429: gap of unknown length
 * 115430 116694: contig of 1265 bp in length
 * 116695 116794: gap of unknown length
 * 116795 118213: contig of 1419 bp in length
 * 118214 118313: gap of unknown length
 * 118314 119716: contig of 1403 bp in length
 * 119717 119816: gap of unknown length
 * 119817 121014: contig of 1198 bp in length
 * 121015 121114: gap of unknown length
 * 121115 122794: contig of 1680 bp in length
 * 122795 122894: gap of unknown length
 * 122895 124962: contig of 2068 bp in length
 * 124963 125062: gap of unknown length
 * 125063 126140: contig of 1078 bp in length
 * 126141 126240: gap of unknown length
 * 126241 127970: contig of 1730 bp in length
 * 127971 128070: gap of unknown length
 * 128071 129564: contig of 1494 bp in length
 * 129565 129664: gap of unknown length
 * 129665 130786: contig of 1122 bp in length
 * 130787 130886: gap of unknown length
 * 130887 132217: contig of 1331 bp in length
 * 132218 132317: gap of unknown length
 * 132318 133604: contig of 1287 bp in length
 * 133605 133704: gap of unknown length
 * 133705 135164: contig of 1460 bp in length
 * 135165 135264: gap of unknown length
 * 135265 136303: contig of 1039 bp in length
 * 136304 136403: gap of unknown length
 * 136404 137437: contig of 1034 bp in length
 * 137438 137537: gap of unknown length
 * 137538 138560: contig of 1023 bp in length
 * 138561 139739: gap of unknown length
 * 139740 139839: contig of 1079 bp in length
 * 139840 141198: contig of 1359 bp in length.

Query Match 1.9%; Score 19; DB 68; Length 141198;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 941 aggggggtaagcgtggtgt 959
 |||||
 Db 129399 AGGGGGTAAGCCTGTGTGT 129381

RESULT 41
 AC084179/c
 LOCUS
 DEFINITION
 AC084179 147009 bp DNA HTG 15-JAN-2001
 Homo sapiens chromosome 5 clone RP11-148P16 map 5, WORKING DRAFT
 SEQUENCE, 39 unordered pieces.
 AC084179
 AC084179.2 GI:12229370
 VERSION
 KEYWORDS
 HTG: HTGS_PHASE1; HTGS_DRAFT.
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 147009)
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 Homo sapiens chromosome 5, clone RP11-148P16
 JOURNAL
 TITLES
 REFERENCES
 2 (bases 1 to 147009)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Beda, F., Boguslavsky, L.,
 Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
 Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, P.,
 Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferrel, P.,
 Fitzhugh, W., Gage, D., Galagan, J., Gardyne, S., Ginde, S., Coyette, M.,
 Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,
 Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Lakoque, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G.,
 Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K.,
 McPheeters, R., Meldrum, J., Meneus, L., Milhova, T., Mienna, V.,
 Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,
 O'Donnell, P., O'Neill, D., Olivari, T.M., Oliver, J., Peterson, K.,
 Pierre, N., Pisanl, C., Pollara, V., Raymond, C., Riebeck, M., Riley, R.,
 Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,
 Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
 Tirrell, A., Travers, M., Trigglio, J., Vassiliev, H., Viel, R., Vo, A.,
 Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zahoun, J.,
 Zimmer, A. and Zody, M.

TITLE
 JOURNAL
 COMMENT
 Submitted (14-Oct-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jan 15, 2001 this sequence version replaced gi:10801407.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seg.wi.mit.edu
 Contact: sequence-submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: 148_P16
 Center clone name: 148_P16
 ----- Summary Statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 132233 bases at least Q40
 Consensus quality: 139576 bases at least Q30
 Consensus quality: 142136 bases at least Q20
 Insert size: 17600; agarose-fp
 Insert size: 143209; sum-of-ctrls
 Quality coverage: 3.0 in Q20 bases; agarose-fp
 Quality coverage: 3.6 in Q20 bases; sum-of-ctrls

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 39 contigs. The true order of the pieces
 * is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1353: contig of 1353 bp in length
1354 1453: gap of 100 bp
1454 2001: contig of 548 bp in length
2002 2101: gap of 100 bp
2102 3124: contig of 1023 bp in length
3125 3224: gap of 100 bp
3225 4546: contig of 1322 bp in length
4547 4646: gap of 100 bp
4647 6015: contig of 1369 bp in length
6016 6115: gap of 100 bp
6116 7259: contig of 1184 bp in length
7300 7399: gap of 100 bp
7400 8836: contig of 1437 bp in length
8837 8936: gap of 100 bp
8937 10282: contig of 1346 bp in length
10283 10382: gap of 100 bp
10383 11352: contig of 970 bp in length
11353 11452: gap of 100 bp
11453 12856: contig of 1404 bp in length
12857 12956: gap of 100 bp
12957 14187: contig of 1231 bp in length
14188 14287: gap of 100 bp
14288 15977: contig of 1690 bp in length
15978 16077: gap of 100 bp
16078 17844: contig of 1767 bp in length
17845 17944: gap of 100 bp
17945 19861: contig of 1917 bp in length
19862 19961: gap of 100 bp
19962 22368: contig of 2407 bp in length
22369 22468: gap of 100 bp
22469 24485: contig of 2017 bp in length
24486 24585: gap of 100 bp
24586 26722: contig of 2137 bp in length
26723 26822: gap of 100 bp
26823 28983: contig of 2161 bp in length
28984 29083: gap of 100 bp
29084 31858: contig of 2775 bp in length
31859 31958: gap of 100 bp
31959 33783: contig of 1825 bp in length
33784 33883: gap of 100 bp
33884 36259: contig of 2376 bp in length
36260 36359: gap of 100 bp
36360 39396: contig of 3037 bp in length
39397 39496: gap of 100 bp
39497 43329: contig of 3833 bp in length
43330 43429: gap of 100 bp
43430 46057: contig of 2628 bp in length
46058 46157: gap of 100 bp
46158 50066: contig of 3909 bp in length
50067 50166: gap of 100 bp
50167 53727: contig of 3561 bp in length
53728 53827: gap of 100 bp
53828 57631: contig of 3804 bp in length
57632 57731: gap of 100 bp
57732 61304: contig of 3573 bp in length
61305 61404: gap of 100 bp
61405 63868: contig of 2464 bp in length
63869 63968: gap of 100 bp
63969 68393: contig of 4425 bp in length
68394 68493: gap of 100 bp
68494 72466: contig of 3973 bp in length
72467 72566: gap of 100 bp
72567 77027: contig of 4461 bp in length
77028 77127: gap of 100 bp
77128 81524: contig of 4397 bp in length
81525 81624: gap of 100 bp
81625 88873: contig of 7249 bp in length
88874 88973: gap of 100 bp

* 88974 96287: contig of 7314 bp in length
* 96288 96387: gap of 100 bp
* 96388 117899: contig of 21512 bp in length
* 117900 117999: gap of 100 bp
* 118000 123663: contig of 5664 bp in length
* 123664 123763: gap of 100 bp
* 123764 137598: contig of 13835 bp in length
* 137599 137698: gap of 100 bp
* 137699 147009: contig of 9311 bp in length.
Location/Qualifiers
1. 147009
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5"
/clone="RP11-148P16"
/clone_id="RPC1-11 Human Male BAC"
1. 1353
/note="assembly-fragment
clone_end:SP6
vector_side:left"
1454. 2001
/note="assembly-fragment"
2102. 3124
/note="assembly-fragment"
3225. 4546
/note="assembly-fragment"
4647. 6015
/note="assembly-fragment"
6116. 7299
/note="assembly-fragment"
7400. 8836
/note="assembly-fragment"
8937. 10282
/note="assembly-fragment"
10383. 11352
/note="assembly-fragment"
11453. 12856
/note="assembly-fragment"
12957. 14187
/note="assembly-fragment"
14288. 15977
/note="assembly-fragment"
16078. 17844
/note="assembly-fragment"
17945. 19861
/note="assembly-fragment"
19962. 22368
/note="assembly-fragment"
22469. 24485
/note="assembly-fragment"
24586. 26722
/note="assembly-fragment"
26823. 28983
/note="assembly-fragment"
29084. 31858
/note="assembly-fragment"
31959. 33783
/note="assembly-fragment"
33884. 36259
/note="assembly-fragment"
36360. 39396

Query Match 1.9%: Score 19; DB 76; Length 147009;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 190 agtagcagctatccaata 208
|||||
Db 10253 AGTAGCAGCTATCCACATA 10235

RESULT 42

AC083840
LOCUS AC083840 156599 bp DNA HTG 03-MAR-2001
DEFINITION Homo sapiens chromosome 8 clone RP11-7319 map 8, WORKING DRAFT
SEQUENCE, 37 unordered pieces.
AC083840
ACCESSION AC083840.2 GI:13184085
VERSION AC083840.2
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 156599)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 8, clone RP11-7319
Unpublished
2 (bases 1 to 156599)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Beda, F., Boguslavsky, L.,
Bouhagiel, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
Chapel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
Deaellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,
FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,
Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,
Iliev, I., Johnson, R., Jones, C., Kann, L., Karas, A., Lakocque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G.,
Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K.,
McNeeters, R., Meldrim, J., Menus, L., Mihova, T., Mieniga, V.,
Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,
O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K.,
Pierre, N., Pisan, C., Pollara, V., Raymond, C., Rieback, M., Riley, R.,
Rogov, P., Rotman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,
Sougnie, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Struss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Travers, M., Triggilo, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, C., Zaitoun, J.,
Zimmer, A. and Zody, M.
Direct Submission
Submitted (03-Oct-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 2, 2001 this sequence version replaced gi:10518398.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L11308
Center clone name: 73_1_9
Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 141695 bases at least Q40
Consensus quality: 148947 bases at least Q30
Consensus quality: 151617 bases at least Q20
Insert size: 152999; sum-of-coverage
Quality coverage: 4.0 in Q20 bases; sum-of-coverage
NOTE: This is a 'working draft' sequence. It currently
consists of 37 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 816: contig of 816 bp in length
* 817 916: gap of 100 bp
* 917 2094: contig of 1178 bp in length
* 2095 2194: gap of 100 bp

2195 3169: contig of 975 bp in length
* 3170 3269: gap of 100 bp
* 3270 3766: contig of 497 bp in length
* 3767 3866: gap of 100 bp
* 3867 5257: contig of 1391 bp in length
* 5258 5357: gap of 100 bp
* 5358 7176: contig of 1819 bp in length
* 7177 7276: gap of 100 bp
* 7277 8366: contig of 1090 bp in length
* 8367 8466: gap of 100 bp
* 8467 9770: contig of 1304 bp in length
* 9771 9870: gap of 100 bp
* 9871 11023: contig of 1153 bp in length
* 11024 11123: gap of 100 bp
* 11124 12210: contig of 1087 bp in length
* 12211 12310: gap of 100 bp
* 12311 13716: contig of 1406 bp in length
* 13717 13816: gap of 100 bp
* 13817 14696: contig of 880 bp in length
* 14697 14796: gap of 100 bp
* 14797 16050: contig of 1254 bp in length
* 16051 16150: gap of 100 bp
* 16151 18691: contig of 2541 bp in length
* 18692 18791: gap of 100 bp
* 18792 20474: contig of 1683 bp in length
* 20475 20574: gap of 100 bp
* 20575 22564: contig of 1990 bp in length
* 22565 22664: gap of 100 bp
* 22665 25349: contig of 2685 bp in length
* 25350 25449: gap of 100 bp
* 25450 28731: contig of 3282 bp in length
* 28732 28831: gap of 100 bp
* 28832 41793: contig of 12662 bp in length
* 41794 41893: gap of 100 bp
* 41894 45562: contig of 3669 bp in length
* 45563 45662: gap of 100 bp
* 45663 49245: contig of 3583 bp in length
* 49246 49345: gap of 100 bp
* 49346 52701: contig of 3356 bp in length
* 52702 52801: gap of 100 bp
* 52802 57587: contig of 4786 bp in length
* 57588 57687: gap of 100 bp
* 57688 60742: contig of 3055 bp in length
* 60743 60842: gap of 100 bp
* 60843 64655: contig of 3813 bp in length
* 64656 64755: gap of 100 bp
* 64756 68878: contig of 4123 bp in length
* 68879 68978: gap of 100 bp
* 68979 73540: contig of 4562 bp in length
* 73541 73640: gap of 100 bp
* 73641 79164: contig of 5524 bp in length
* 79165 79264: gap of 100 bp
* 79265 87667: contig of 8403 bp in length
* 87668 87767: gap of 100 bp
* 87768 96110: contig of 8343 bp in length
* 96111 96210: gap of 100 bp
* 96211 104171: contig of 7961 bp in length
* 104172 104271: gap of 100 bp
* 104272 110622: contig of 6351 bp in length
* 110623 110722: gap of 100 bp
* 110723 121728: contig of 11006 bp in length
* 121729 121828: gap of 100 bp
* 121829 129949: contig of 8121 bp in length
* 129950 130049: gap of 100 bp
* 130050 139747: contig of 9688 bp in length
* 139748 139847: gap of 100 bp
* 139848 153810: contig of 13863 bp in length
* 153811 153910: gap of 100 bp
* 153911 156599: contig of 2689 bp in length.
Location/Qualifiers
1. 156599
/organism="Homo sapiens"
/db_xref="taxon:9606"


```
/chromosome="8"
/map="8"
/clone_lib="RP11-7319"
/clone_lib="RP11-11 Human Male BAC"
1. 816
/misc_feature /note="assembly_fragment"
vector_side:left"
917. 2094
/misc_feature /note="assembly_fragment"
2195. 3169
/misc_feature /note="assembly_fragment"
3270. 3766
/misc_feature /note="assembly_fragment"
3867. 5257
/misc_feature /note="assembly_fragment"
5358. 7176
/misc_feature /note="assembly_fragment"
7277. 8366
/misc_feature /note="assembly_fragment"
8467. 9770
/misc_feature /note="assembly_fragment"
9871. 11023
/misc_feature /note="assembly_fragment"
11124. 12210
/misc_feature /note="assembly_fragment"
12311. 13716
/misc_feature /note="assembly_fragment"
13817. 14696
/misc_feature /note="assembly_fragment"
14797. 16050
/misc_feature /note="assembly_fragment"
16151. 18691
/misc_feature /note="assembly_fragment"
18792. 20474
/misc_feature /note="assembly_fragment"
20575. 22564
/misc_feature /note="assembly_fragment"
22665. 25349
/misc_feature /note="assembly_fragment"
25450. 28731
/misc_feature /note="assembly_fragment"
28832. 41793
/misc_feature /note="assembly_fragment"
41894. 45562
/misc_feature /note="assembly_fragment"
45663. 49245
/misc_feature /note="assembly_fragment"
49346. 52701
/misc_feature /note="assembly_fragment"
52802. 57587
/misc_feature /note="assembly_fragment"
57688. 60742
/misc_feature /note="assembly_fragment"
60843. 64655

Query Match 1.98; Score 19; DB 76; Length 156599;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 243 cccgttcagccagcc 261
|||||
Db 2492 CCCGTTTCAGCCGCGCACC 2510
```

```
RESULT 43
AC034236/c AC034236 160404 bp DNA HTG 18-JUL-2000
LOCUS Homo sapiens chromosome 5 clone CTD-2287016, WORKING DRAFT
DEFINITION
SEQUENCE 16 ordered pieces.
ACCESSION AC034236
VERSION AC034236.2 GI:9256733
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
```

```
SOURCE
ORGANISM human.
REFERENCE
AUTHORS 1 (bases 1 to 160404)
TITLE DOE Joint Genome Institute.
JOURNAL Sequencing of Human Chromosome 5
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 160404)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 18, 2000 this sequence version replaced gi:7417713.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 738109
Center clone name: CITB-RL-2287016
-----
Summary Statistics
Consensus quality: 148274 bases at least Q40
Consensus quality: 155685 bases at least Q30
Consensus quality: 157705 bases at least Q20
Estimated insert size: 145000; pulse field gel estimation
Estimated insert size: 159704; sum-of-contrigs estimation
Quality coverage: 6.47 in Q20 bases; pulse field gel estimation
Quality coverage: 5.87 in Q20 bases; sum-of-contrigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contrigs. Gaps between the contrigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 16110: contig of 16110 bp in length
1 16111 16210: gap of unknown length
1 16211 38617: contig of 22407 bp in length
1 38618 38717: gap of unknown length
1 38718 71154: contig of 32437 bp in length
1 71155 71254: gap of unknown length
1 71255 73939: contig of 2685 bp in length
1 73940 74039: gap of unknown length
1 74040 75830: contig of 1791 bp in length
1 75831 75930: gap of unknown length
1 75931 79883: contig of 3953 bp in length
1 79884 79983: gap of unknown length
1 79984 98933: contig of 18950 bp in length
1 98934 99033: gap of unknown length
1 99034 106554: contig of 7521 bp in length
1 106555 106654: gap of unknown length
1 106655 112876: contig of 6222 bp in length
1 112877 112976: gap of unknown length
1 112977 115635: contig of 2659 bp in length
1 115636 115735: gap of unknown length
1 115736 132099: contig of 16364 bp in length
1 132100 132199: gap of unknown length
1 132200 138242: contig of 6043 bp in length
1 138243 138342: gap of unknown length
1 138343 140816: contig of 2474 bp in length
1 140817 140916: gap of unknown length
1 140917 145928: contig of 5012 bp in length
1 145929 146028: gap of unknown length
1 146029 156369: contig of 10341 bp in length
1 156370 156469: gap of unknown length
1 156470 160404: contig of 3935 bp in length.
Location/Qualifiers
1. 160404
```

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone_lib="Caltech human BAC library D"
 BASE COUNT 46000 a 28337 c 31268 g 53298 t 1501 others
 ORIGIN

Query Match 1.9% Score 19: DB 71: Length 160404;
 Best Local Similarity 100.0% Pred. No. 33;
 Matches 19: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Ox 800 tagagcagacagatgcaga 818
 |||||
 DB 141561 TAGAGCAGACAGATGCAGA 141543

RESULT 44
 AC015815/C

LOCUS AC015815 161549 bp DNA HTG 30-MAR-2000
 DEFINITION Homo sapiens clone RP11-21015, WORKING DRAFT SEQUENCE, 21 unordered

ACCESSION

AC015815

VERSION AC015815.4 GI:7341971

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS

Blumen, B., Linton, L., Nusbaum, C. and Lander, E.

TITLE

Homo sapiens, clone RP11-21015

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 161549)

AUTHORS

Blumen, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,

TITLE

Baldwin, J., Barna, N., Beckert, R., Boguslavsky, L., Boukhallal, B.,

JOURNAL

Brown, A., Castle, A., Colangelo, M., Collins, S., Collamore, A.,

AUTHORS

Cooke, P., DeArnell, K., Dewar, K., Domino, M., Donnell, L., Doyle, M.,

TITLE

Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,

JOURNAL

Gallagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,

AUTHORS

Howland, J. C., Johnson, R., Jones, C., Kann, L., Karates, A., Klein, J.,

TITLE

Lehoczky, D., Lien, C., Locke, K., MacDonald, P., Marquis, N.,

JOURNAL

McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Melnick, J.,

AUTHORS

Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,

TITLE

Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,

JOURNAL

Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,

AUTHORS

Testa, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,

TITLE

Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

JOURNAL

Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome

COMMENT

Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 30, 2000 this sequence version replaced gi:6939989.
 ALL repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE

Center: Whitehead Institute/ MIT Center for Genome Research

JOURNAL

Center: Whitehead Institute/ MIT Center for Genome Research

COMMENT

Web site: http://www-seq.wi.mit.edu

TITLE

Contact: sequence_submissions@genome.wi.mit.edu

JOURNAL

Project Information

COMMENT

Center project name: 14073

TITLE

Center clone name: 21_O15

JOURNAL

Summary Statistics

COMMENT

Sequencing vector: M13: M77815: 100% of reads

TITLE

Chemistry: Dye-terminator Big Dye 100% of reads

JOURNAL

Assembly program: Phrap; version 0.960731

COMMENT

Consensus quality: 15133 bases at least Q40

TITLE

Consensus quality: 15764 bases at least Q30

JOURNAL

Consensus quality: 157650 bases at least Q20

COMMENT

Insert size: 15700; agarose-1p

Insert size: 159549; sum-of-ctrls

Quality coverage: 4.9 in Q20 bases; agarose-1p
 Quality coverage: 4.8 in Q20 bases; sum-of-ctrls

NOTE: This is a 'working draft' sequence. It currently
 consists of 21 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 1138: contig of 1138 bp in length
 1139 1238: gap of 100 bp
 1239 2713: contig of 1475 bp in length
 2714 2813: gap of 100 bp
 2814 4948: contig of 2135 bp in length
 4949 5048: gap of 100 bp
 5049 7939: contig of 2891 bp in length
 7940 8039: gap of 100 bp
 8040 9902: contig of 1863 bp in length
 9903 10002: gap of 100 bp
 10003 12571: contig of 2569 bp in length
 12572 12671: gap of 100 bp
 12672 16792: contig of 4121 bp in length
 16793 16892: gap of 100 bp
 16893 21320: contig of 4428 bp in length
 21321 21420: gap of 100 bp
 21421 28365: contig of 6945 bp in length
 28366 28465: gap of 100 bp
 28466 34699: contig of 6234 bp in length
 34700 34799: gap of 100 bp
 34800 40251: contig of 5452 bp in length
 40252 40351: gap of 100 bp
 40352 47897: contig of 7546 bp in length
 47898 47997: gap of 100 bp
 47998 55766: contig of 7769 bp in length
 55767 55866: gap of 100 bp
 55867 63968: contig of 8102 bp in length
 63969 64068: gap of 100 bp
 64069 74694: contig of 10626 bp in length
 74695 74794: gap of 100 bp
 74795 84737: contig of 9943 bp in length
 84738 84837: gap of 100 bp
 84838 98320: contig of 13483 bp in length
 98321 98420: gap of 100 bp
 98421 110611: contig of 12191 bp in length
 110612 110711: gap of 100 bp
 110712 122418: contig of 11707 bp in length
 122419 122518: gap of 100 bp
 122519 137004: contig of 14486 bp in length
 137005 137104: gap of 100 bp
 137105 161549: contig of 24445 bp in length.

FEATURES
 source

1. 161549
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="RP11-21015"
 /clone_lib="RP11-21015"
 1. 1138
 /note="assembly-fragment"
 misc_feature 1239..2713
 /note="assembly-fragment"
 misc_feature 2814..4948
 /note="assembly-fragment"
 misc_feature 5049..7939
 /note="assembly-fragment"
 misc_feature 8040..9902
 /note="assembly-fragment"
 misc_feature 10003..12571
 /note="assembly-fragment"
 misc_feature 12672..16792
 /note="assembly-fragment"
 misc_feature 16893..21320

misc_feature /note="assembly_fragment"
21421..28365
/note="assembly_fragment"
misc_feature 28466..33469
/note="assembly_fragment"
clone_end:SP6
vector_side:left
misc_feature 34800..40251
/note="assembly_fragment"
misc_feature 40352..47897
/note="assembly_fragment"
misc_feature 47998..55766
/note="assembly_fragment"
clone_end:T7
vector_side:left
misc_feature 55867..63968
/note="assembly_fragment"
misc_feature 64069..74694
/note="assembly_fragment"
misc_feature 74795..84737
/note="assembly_fragment"
misc_feature 84838..98320
/note="assembly_fragment"
misc_feature 98421..110611
/note="assembly_fragment"
misc_feature 110712..122418
/note="assembly_fragment"
misc_feature 122519..137004
/note="assembly_fragment"
misc_feature 137105..161549
/note="assembly_fragment"
BASE COUNT 41638 a 38819 c 39363 g 39720 t 2009 others
ORIGIN

Query Match 1.9%; Score 19; DB 64; Length 161549;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 105 cagggaagaacagcagcca 123
|||||
Db 28267 CAGGAGAACACGACGCA 28249

RESULT 45
AC009108 AC009108 170431 bp DNA HTG 02-SEP-2000
LOCUS Homo sapiens chromosome 16 clone RP11-46309, WORKING DRAFT
DEFINITION
SEQUENCE, 4 ordered pieces.
AC009108
AC009108.8 GI:9964740
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 170431)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 170431)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-APR-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Sep 2, 2000 this sequence version replaced gi:8575957.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web Site: http://www.jgi.doe.gov

Project Information
Center Project Name: 584502

Center clone name: RPCI-11_46309

Summary Statistics
Consensus quality: 167345 bases at least Q40
Consensus quality: 169423 bases at least Q30
Consensus quality: 169900 bases at least Q20
Estimated insert size: 181440; agarose-fp estimation
Estimated insert size: 170331; sum-of-ctdgs estimation
Quality coverage: 10.63 in Q20 bases; agarose-fp estimation
Quality coverage: 11.33 in Q20 bases; sum-of-ctdgs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 24820: contig of 24820 bp in length
* 24821 24920: gap of unknown length
* 24921 37862: contig of 12942 bp in length
* 37863 37962: gap of unknown length
* 37963 114671: contig of 76709 bp in length
* 114672 114771: gap of unknown length
* 114772 170431: contig of 55660 bp in length.
FEATURES
Location/Qualifiers
1..170431
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-46309"
/clone_lib="RPCI human BAC library 11"
BASE COUNT 41117 a 43130 c 42591 g 43290 t 303 others
ORIGIN

Query Match 1.9%; Score 19; DB 61; Length 170431;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 280 gccgcgcgcgcgcgcgcgcgc 298
|||||
Db 21976 GCCGCGCGCGCATCGCGTG 21994

RESULT 46
AC013343 AC013343 175249 bp DNA HTG 30-MAR-2000
LOCUS Homo sapiens clone RP11-22B10, WORKING DRAFT SEQUENCE, 13 unordered
DEFINITION
pieces.
AC013343
AC013343.3 GI:7341972
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 175249)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barina, N., Beckert, R., Boguslavsky, L., Bouckgalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., DeRellano, K., Dewar, K., Domingo, M., Donelan, L., Doyle, M.,
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J.,
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoczky, J., Liu, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,

TITLE
JOURNAL
COMMENT

Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testfay, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wymann, D., Ye, W. J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (06-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 30, 2000 this sequence version replaced g1:6478975.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www.seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center project name: 14088
Center clone name: 22_B_10

Summary Statistics
Sequencing vector: M13: M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 169195 bases at least Q40
Consensus quality: 171919 bases at least Q30
Consensus quality: 172990 bases at least Q20
Insert size: 172000; agarose-fp
Insert size: 174049; sum-of-ctrls
Quality coverage: 6.8 in Q20 bases; agarose-fp
Quality coverage: 6.7 in Q20 bases; sum-of-ctrls

NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1100: contig of 1100 bp in length
* 1101 1200: gap of 100 bp
* 1201 2923: contig of 1723 bp in length
* 2924 3023: gap of 100 bp
* 3024 6205: contig of 3182 bp in length
* 6206 6305: gap of 100 bp
* 6306 8632: contig of 2327 bp in length
* 8633 8732: gap of 100 bp
* 8733 16009: contig of 7277 bp in length
* 16010 16109: gap of 100 bp
* 16110 26909: contig of 10800 bp in length
* 26910 27009: gap of 100 bp
* 27010 39473: contig of 12464 bp in length
* 39474 39573: gap of 100 bp
* 39574 51751: contig of 12178 bp in length
* 51752 51851: gap of 100 bp
* 51852 66815: contig of 14964 bp in length
* 66816 66915: gap of 100 bp
* 66916 79391: contig of 12476 bp in length
* 79392 79491: gap of 100 bp
* 79492 105904: contig of 26413 bp in length
* 105905 106004: gap of 100 bp
* 106005 130201: contig of 24197 bp in length
* 130202 130301: gap of 100 bp
* 130302 175249: contig of 44948 bp in length.
Location/Qualifiers
1. 175249
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-22B10"
/clone_1fb="RP11-11 Human Male BAC"
1. 1100
/note="assembly_fragment"
misc_feature
1201. 2923

misc_feature
/note="assembly_fragment"
3024. 6205
/note="assembly_fragment"
misc_feature
6306. 8632
/note="assembly_fragment"
clone_end:77
vector_side:right"
7933. 16009
/note="assembly_fragment"
16110. 26909
/note="assembly_fragment"
27010. 39473
/note="assembly_fragment"
39574. 51751
/note="assembly_fragment"
51852. 66815
/note="assembly_fragment"
66916. 79391
/note="assembly_fragment"
79492. 105904
/note="assembly_fragment"
106005. 130201
/note="assembly_fragment"
130302. 175249
/note="assembly_fragment"
clone_end:SP6
vector_side:left"

BASE COUNT 55507 a 34489 c 33847 g 50204 t 1202 others
ORIGIN

Query Match 1.9%; Score 19; DB 63; Length 175249;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 agtagcagtcataccaata 208
Db 68509 AGTAGCAGTCATACCAATA 68491
IIIIIIIIIIIIIIIIIIII

RESULT 47
AL356772
LOCUS
DEFINITION
Homo sapiens chromosome 10 clone RP11-47B24, *** SEQUENCING IN
PROGRESS ***; 8 unordered pieces.
ACCESSION
AL356772
VERSION
AL356772.7 GI:13624524
KEYWORDS
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Granulata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 178751)
REFERENCE
Burton, J.
Direct Submission
Submitted (11-APR-2001) Sanger Centre, Hinxton, Cambridgehire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Apr 14, 2001 this sequence version replaced g1:13620380.
COMMENT
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: BA47B24
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 174140 bases at least Q40
Consensus quality: 175949 bases at least Q30
Consensus quality: 177051 bases at least Q20

Insert size: 178051; sum-of-contigs
Insert size: 150047; 16.7% error; agarose-fp
Quality coverage: 5.40x in Q20 bases; sum-of-contigs Quality
coverage: 6.47x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2302: contig of 2302 bp in length
* 2303 2402: gap of 100 bp
* 2403 28561: contig of 26159 bp in length
* 28562 28661: gap of 100 bp
* 28662 31199: contig of 2538 bp in length
* 31200 31299: gap of 100 bp
* 31300 33428: contig of 2129 bp in length
* 33429 33528: gap of 100 bp
* 33529 41182: contig of 7654 bp in length
* 41183 41282: gap of 100 bp
* 41283 98057: contig of 56775 bp in length
* 98058 98157: gap of 100 bp
* 98158 104236: contig of 6079 bp in length
* 104237 104336: gap of 100 bp
* 104337 178751: contig of 74415 bp in length.
Location/Qualifiers
1. 178751

FEATURES

SOURCE

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone_id="RP11-47B24"
1. 2302
/note="assembly_fragment:02043
fragment_chain:1"
2403. 28561
/note="assembly_fragment:00161
fragment_chain:1"
28662. 31199
/note="assembly_fragment:02019
fragment_chain:1"
31300. 33428
/note="assembly_fragment:01183"
33529. 41182
/note="assembly_fragment:01813"
41283. 98057
/note="assembly_fragment:01877"
98158. 104236
/note="assembly_fragment:02146"
104337. 178751
/note="assembly_fragment:02387"
BASE COUNT 53832 a 34081 c 34654 g 55484 t 700 others
ORIGIN

Query Match 1.9% Score 19; DB 80; Length 178751;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 807 acagatcacagactgcag 825
|||||
Db 96823 ACAGATCACAGACTGCAG 96841

RESULT 48
AP002501/c DNA HTG 13-JUN-2000
LOCUS Homo sapiens chromosome 18 clone RP11-680K13 map 18q12, WORKING
DEFINITION DRAFT SEQUENCE, 26 unordered pieces.
ACCESSION AP002501

VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 179844)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE
JOURNAL
PUBLISHED ONLY IN DATABASE (2000) In press
2 (bases 1 to 179844)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

AP002501.1 GI:8547586
HTG: HTGS, PHASE1; HTGS: DRAFT.
Homo sapiens DNA, clone:RP11-680K13.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 179844)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 179,844 genomic DNA of 18q12
Published Only in Database (2000) In press
2 (bases 1 to 179844)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (09-JUN-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail:hattori@gsr.riken.go.jp,
URL:http://bhp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://bhp.gsc.riken.go.jp/
Contact: hattori@gsr.riken.go.jp
----- Project Information
Center project name: HumPrat18
Center clone name: RP11-680K13
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 15359 bases at least Q40
Consensus quality: 16513 bases at least Q30
Consensus quality: 173786 bases at least Q20
Insert size: 177344; sum-of-contigs
Quality coverage: 4.16x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
26 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved

1 18894 contig of 18894 bp in length
19995 39625 contig of 19631 bp in length
39726 54530 contig of 14805 bp in length
54631 68239 contig of 13609 bp in length
68340 78114 contig of 9775 bp in length
78215 89873 contig of 11659 bp in length
89974 99848 contig of 9875 bp in length
99949 109211 contig of 9283 bp in length
109312 116907 contig of 7566 bp in length
117008 121661 contig of 4654 bp in length
121762 128613 contig of 6852 bp in length
128714 132862 contig of 4149 bp in length
132963 137321 contig of 4359 bp in length
137422 142209 contig of 4788 bp in length
142310 146753 contig of 4444 bp in length
146854 152208 contig of 5355 bp in length
152309 158467 contig of 6139 bp in length
158568 162342 contig of 3975 bp in length
162643 165710 contig of 3068 bp in length
165811 168812 contig of 3002 bp in length
168913 170572 contig of 1660 bp in length
170673 172454 contig of 1782 bp in length
172555 175531 contig of 2977 bp in length
175632 177476 contig of 1845 bp in length
177577 178671 contig of 1095 bp in length
178772 179844 contig of 1073 bp in length.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 26 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 19894: contig of 19894 bp in length
19895 19994: gap of 100 bp
19995 39625: contig of 19631 bp in length
39626 39725: gap of 100 bp
39726 54530: contig of 14805 bp in length
54531 54630: gap of 100 bp
54631 68239: contig of 13609 bp in length
68240 68339: gap of 100 bp
68340 78114: contig of 9775 bp in length
78115 78214: gap of 100 bp
78215 89873: contig of 11659 bp in length
89874 89973: gap of 100 bp
89974 99848: contig of 9875 bp in length
99849 99948: gap of 100 bp
99949 109211: contig of 9263 bp in length
109212 109311: gap of 100 bp
109312 116907: contig of 7596 bp in length
116908 117007: gap of 100 bp
117008 121661: contig of 4654 bp in length
121662 121761: gap of 100 bp
121762 128613: contig of 6852 bp in length
128614 128713: gap of 100 bp
128714 132862: contig of 4149 bp in length
132863 132962: gap of 100 bp
132963 137321: contig of 4359 bp in length
137322 137421: gap of 100 bp
137422 142209: contig of 4788 bp in length
142210 142309: gap of 100 bp
142310 146753: contig of 4444 bp in length
146754 146853: gap of 100 bp
146854 152208: contig of 5355 bp in length
152209 152308: gap of 100 bp
152309 158467: contig of 6159 bp in length
158468 158567: gap of 100 bp
158568 162542: contig of 3975 bp in length
162543 162642: gap of 100 bp
162643 165710: contig of 3068 bp in length
165711 165810: gap of 100 bp
165811 168812: contig of 3002 bp in length
168813 168912: gap of 100 bp
168913 170572: contig of 1660 bp in length
170573 170672: gap of 100 bp
170673 172454: contig of 1782 bp in length
172455 172554: gap of 100 bp
172555 175531: contig of 2977 bp in length
175532 175631: gap of 100 bp
175632 177476: contig of 1845 bp in length
177477 177576: gap of 100 bp
177577 178671: contig of 1095 bp in length
178672 178771: gap of 100 bp
178772 179844: contig of 1073 bp in length.
Location/Qualifiers
1. 179844
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18q12"
/clone="RP11-680K13"
1. 19894
/misc_feature
19955 19994: gap of 100 bp
/misc_feature
19995 39625: contig of 19631 bp in length
/misc_feature
39726 39725: gap of 100 bp
/misc_feature
54531 54630: gap of 100 bp
/misc_feature
54631 68239: contig of 13609 bp in length
/misc_feature
68240 68339: gap of 100 bp
/misc_feature
68340 78114: contig of 9775 bp in length
/misc_feature
78115 78214: gap of 100 bp
/misc_feature
78215 89873: contig of 11659 bp in length
/misc_feature
89874 89973: gap of 100 bp
/misc_feature
89974 99848: contig of 9875 bp in length
/misc_feature
99849 99948: gap of 100 bp
/misc_feature
99949 109211: contig of 9263 bp in length
/misc_feature
109212 109311: gap of 100 bp
/misc_feature
109312 116907: contig of 7596 bp in length
/misc_feature
116908 117007: gap of 100 bp
/misc_feature
117008 121661: contig of 4654 bp in length
/misc_feature
121662 121761: gap of 100 bp
/misc_feature
121762 128613: contig of 6852 bp in length
/misc_feature
128614 128713: gap of 100 bp
/misc_feature
128714 132862: contig of 4149 bp in length
/misc_feature
132863 132962: gap of 100 bp
/misc_feature
132963 137321: contig of 4359 bp in length
/misc_feature
137322 137421: gap of 100 bp
/misc_feature
137422 142209: contig of 4788 bp in length
/misc_feature
142210 142309: gap of 100 bp
/misc_feature
142310 146753: contig of 4444 bp in length
/misc_feature
146754 146853: gap of 100 bp
/misc_feature
146854 152208: contig of 5355 bp in length
/misc_feature
152209 152308: gap of 100 bp
/misc_feature
152309 158467: contig of 6159 bp in length
/misc_feature
158468 158567: gap of 100 bp
/misc_feature
158568 162542: contig of 3975 bp in length
/misc_feature
162543 162642: gap of 100 bp
/misc_feature
162643 165710: contig of 3068 bp in length
/misc_feature
165711 165810: gap of 100 bp
/misc_feature
165811 168812: contig of 3002 bp in length
/misc_feature
168813 168912: gap of 100 bp
/misc_feature
168913 170572: contig of 1660 bp in length
/misc_feature
170573 170672: gap of 100 bp
/misc_feature
170673 172454: contig of 1782 bp in length
/misc_feature
172455 172554: gap of 100 bp
/misc_feature
172555 175531: contig of 2977 bp in length
/misc_feature
175532 175631: gap of 100 bp
/misc_feature
175632 177476: contig of 1845 bp in length
/misc_feature
177477 177576: gap of 100 bp
/misc_feature
177577 178671: contig of 1095 bp in length
/misc_feature
178672 178771: gap of 100 bp
/misc_feature
178772 179844: contig of 1073 bp in length.

```

```

FEATURES
source
1. 179844
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18q12"
/clone="RP11-680K13"
1. 19894
/misc_feature
19955 19994: gap of 100 bp
/misc_feature
19995 39625: contig of 19631 bp in length
/misc_feature
39726 39725: gap of 100 bp
/misc_feature
54531 54630: gap of 100 bp
/misc_feature
54631 68239: contig of 13609 bp in length
/misc_feature
68240 68339: gap of 100 bp
/misc_feature
68340 78114: contig of 9775 bp in length
/misc_feature
78115 78214: gap of 100 bp
/misc_feature
78215 89873: contig of 11659 bp in length
/misc_feature
89874 89973: gap of 100 bp
/misc_feature
89974 99848: contig of 9875 bp in length
/misc_feature
99849 99948: gap of 100 bp
/misc_feature
99949 109211: contig of 9263 bp in length
/misc_feature
109212 109311: gap of 100 bp
/misc_feature
109312 116907: contig of 7596 bp in length
/misc_feature
116908 117007: gap of 100 bp
/misc_feature
117008 121661: contig of 4654 bp in length
/misc_feature
121662 121761: gap of 100 bp
/misc_feature
121762 128613: contig of 6852 bp in length
/misc_feature
128614 128713: gap of 100 bp
/misc_feature
128714 132862: contig of 4149 bp in length
/misc_feature
132863 132962: gap of 100 bp
/misc_feature
132963 137321: contig of 4359 bp in length
/misc_feature
137322 137421: gap of 100 bp
/misc_feature
137422 142209: contig of 4788 bp in length
/misc_feature
142210 142309: gap of 100 bp
/misc_feature
142310 146753: contig of 4444 bp in length
/misc_feature
146754 146853: gap of 100 bp
/misc_feature
146854 152208: contig of 5355 bp in length
/misc_feature
152209 152308: gap of 100 bp
/misc_feature
152309 158467: contig of 6159 bp in length
/misc_feature
158468 158567: gap of 100 bp
/misc_feature
158568 162542: contig of 3975 bp in length
/misc_feature
162543 162642: gap of 100 bp
/misc_feature
162643 165710: contig of 3068 bp in length
/misc_feature
165711 165810: gap of 100 bp
/misc_feature
165811 168812: contig of 3002 bp in length
/misc_feature
168813 168912: gap of 100 bp
/misc_feature
168913 170572: contig of 1660 bp in length
/misc_feature
170573 170672: gap of 100 bp
/misc_feature
170673 172454: contig of 1782 bp in length
/misc_feature
172455 172554: gap of 100 bp
/misc_feature
172555 175531: contig of 2977 bp in length
/misc_feature
175532 175631: gap of 100 bp
/misc_feature
175632 177476: contig of 1845 bp in length
/misc_feature
177477 177576: gap of 100 bp
/misc_feature
177577 178671: contig of 1095 bp in length
/misc_feature
178672 178771: gap of 100 bp
/misc_feature
178772 179844: contig of 1073 bp in length.

```

```

misc_feature 68340..78114
             /note="assembly_fragment"
misc_feature 78215..89873
             /note="assembly_fragment"
misc_feature 89974..99848
             /note="assembly_fragment"
misc_feature 99949..109211
             /note="assembly_fragment"
misc_feature 109312..116907
             /note="assembly_fragment"
misc_feature 117008..121661
             /note="assembly_fragment"
misc_feature 121762..128613
             /note="assembly_fragment"
misc_feature 128714..132862
             /note="assembly_fragment"
misc_feature 132963..137321
             /note="assembly_fragment"
misc_feature 137422..142209
             /note="assembly_fragment"
misc_feature 142310..146753
             /note="assembly_fragment"
misc_feature 146854..152208
             /note="assembly_fragment"
misc_feature 152309..158467
             /note="assembly_fragment"
misc_feature 158568..162542
             /note="assembly_fragment"
misc_feature 162643..165710
             /note="assembly_fragment"
misc_feature 165811..168812
             /note="assembly_fragment"
misc_feature 168913..170572
             /note="assembly_fragment"
misc_feature 170673..172454
             /note="assembly_fragment"
misc_feature 172535..175531
             /note="assembly_fragment"
misc_feature 175632..177476
             /note="assembly_fragment"
misc_feature 177577..178671
             /note="assembly_fragment"
misc_feature 178772..179844
             /note="assembly_fragment"
misc_feature 179845..180000
             /note="assembly_fragment"
BASE COUNT 42717 a 45958 c 47394 g 41275 t 2500 others
ORIGIN
Query Match 1.9% Score 19; DB 83; Length 179844;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 243 ccctgttcagccagcc 261
Db 99697 CCCTGTTTCAGCCAGCC 99679
RESULT 49
AC068169/c 180717 bp DNA HTG 15-NOV-2000
LOCUS Homo sapiens chromosome 17 clone RP11-651B2 map 17, WORKING DRAFT
DEFINITION SEQUENCE, 15 unordered pieces.
ACCESSION AC068169
VERSION AC068169.3 GI:1136767
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 180717)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone RP11-651B2
JOURNAL Unpublished

```

REFERENCE
AUTHORS

2 (bases 1 to 180717)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F., Boguslavsky, L., Bonkhgalter, B., Brown, A., Burkett, G., Campiolo, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J., Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyle, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J., Laroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGuck, A., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mianga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testave, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (29-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Nov 10, 2000 this sequence version replaced gi:8389515.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WtBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L10237
 Center clone name: 651_B-2

----- Summary Statistics
 Sequencing vector: M13; M7815; 43% of reads
 Sequencing vector: Plasmid; n/a; 57% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 173839 bases at least Q40
 Consensus quality: 177110 bases at least Q30
 Consensus quality: 178528 bases at least Q20
 Insert size: 179000; agarose-fp
 Insert size: 179317; sum-of-ctrls
 Quality coverage: 8.5 in Q20 b.
 Quality coverage: 8.5 in Q20 b.
 NOTE: This is a 'working draft' sequence. It currently
 * consists of 15 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1
 757 856: contig of 756 bp in length
 857 2208: contig of 1352 bp in length
 2209 2308: gap of 100 bp
 2309 3715: contig of 1407 bp in length
 3716 3815: gap of 100 bp
 3816 5977: contig of 2162 bp in length
 5978 6077: gap of 100 bp
 6078 57570: contig of 51493 bp in length
 57571 57670: gap of 100 bp
 57671 60754: contig of 3084 bp in length
 60755 60854: gap of 100 bp
 60855 65276: contig of 4422 bp in length
 65277 65376: gap of 100 bp
 65377 71222: contig of 5846 bp in length
 71223 71322: gap of 100 bp

71323 81741: contig of 10419 bp in length
 * 81742 81841: gap of 100 bp
 * 81842 91954: contig of 10113 bp in length
 * 91955 92054: gap of 100 bp
 * 92055 104839: contig of 12785 bp in length
 * 104840 104939: gap of 100 bp
 * 104940 125196: contig of 20257 bp in length
 * 125197 125296: gap of 100 bp
 * 125297 149973: contig of 24677 bp in length
 * 149974 150073: gap of 100 bp
 * 150074 178960: contig of 28887 bp in length
 * 178961 179060: gap of 100 bp
 * 179061 180717: contig of 1657 bp in length.

FEATURES
 source
 1..180717
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17"
 /clone="RP11-651B2"
 /clone_id="RP11-11 Human Male BAC"
 1..756
 /note="assembly-fragment
 clone_end:SP6
 vector_side:left"
 857..2208
 /note="assembly-fragment"
 2309..3715
 /note="assembly-fragment"
 3816..5977
 /note="assembly-fragment"
 6078..57570
 /note="assembly-fragment"
 57671..60754
 /note="assembly-fragment"
 60855..65276
 /note="assembly-fragment"
 65377..71222
 /note="assembly-fragment"
 71323..81741
 /note="assembly-fragment"
 81842..91954
 /note="assembly-fragment"
 92055..104839
 /note="assembly-fragment"
 104940..125196
 /note="assembly-fragment"
 125297..149973
 /note="assembly-fragment"
 150074..178960
 /note="assembly-fragment"
 179061..180717
 /note="assembly-fragment
 clone_end:T7
 vector_side:right"

BASE COUNT 48215 a 43298 c 41795 g 46002 t 1407 others
 ORIGIN

Query Match 1.9%: Score 19; DB 73; Length 180717;
 Best Local Similarity 100.0%: Pred. No. 32;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 105 cagggaagacagcagca 123
 |||||
 Db 132375 CAGGGAAGACAGCAGCA 132357

RESULT 50
 AC024507 181853 bp DNA HTG 16-MAR-2000
 LOCUS Homo sapiens chromosome 1 clone RP11-153J19 map 1, *** SEQUENCING
 DEFINITION IN PROGRESS ***, 55 unordered pieces.

ACCESSION AC024507
VERSION AC024507.2 GI:7249351
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 181853)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 1, clone RP11-153J19
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 181853)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavsky,L.,
Boukhalter,B., Brown,A., Burkett,G., Campoliano,A., Castle,A.,
Chapel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
Deaellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M.,
Fenster,J., Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D.,
Galgan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Landers,T., Largocque,K., Lehoczy,J., Levine,R.,
Lieu,C., Liu,G., Locke,K., Macdonald,P., Margulis,N., McCarthy,M.,
McEwan,P., McQuirk,A., McKernan,K., McPheters,R., Meldrum,J.,
Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Ollivar,T.M.,
Peterson,K., Pierre,N., Pisanil,C., Pollara,V., Raymond,C.,
Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
Severy,P., Spencer,B., Stange-Thoman,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Titrill,A.,
Travers,M., Triggillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
Zody,M.
Direct Submission
Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 16, 2000 this sequence version replaced gi:7108303.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7202
Center clone name: 153_J19

* NOTE: This is a 'working draft' sequence. It currently
* consists of 55 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 1225: contig of 1225 bp in length
* 1226 1325: gap of 100 bp
* 1326 2383: contig of 1058 bp in length
* 2384 2483: gap of 100 bp
* 2484 3638: contig of 1155 bp in length
* 3639 3738: gap of 100 bp
* 3739 4871: contig of 1133 bp in length
* 4872 4971: gap of 100 bp
* 4972 6453: contig of 1482 bp in length
* 6454 6553: gap of 100 bp
* 6554 7651: contig of 1098 bp in length
* 7652 7751: gap of 100 bp
* 7752 8917: contig of 1166 bp in length
* 8918 9017: gap of 100 bp
* 9018 10865: contig of 1848 bp in length
* 10866 10965: gap of 100 bp

10966 12050: contig of 1085 bp in length
* 12051 12150: gap of 100 bp
* 12151 13291: contig of 1141 bp in length
* 13292 13391: gap of 100 bp
* 13392 14491: contig of 1100 bp in length
* 14492 14591: gap of 100 bp
* 14592 15649: contig of 1058 bp in length
* 15650 15749: gap of 100 bp
* 15750 17150: contig of 1401 bp in length
* 17151 17250: gap of 100 bp
* 17251 18421: contig of 1171 bp in length
* 18422 18521: gap of 100 bp
* 18522 20045: contig of 1524 bp in length
* 20046 20145: gap of 100 bp
* 20146 21485: contig of 1340 bp in length
* 21486 21585: gap of 100 bp
* 21586 23169: contig of 1584 bp in length
* 23170 23269: gap of 100 bp
* 23270 24275: contig of 1006 bp in length
* 24276 24375: gap of 100 bp
* 24376 25509: contig of 1134 bp in length
* 25510 25609: gap of 100 bp
* 25610 26831: contig of 1222 bp in length
* 26832 26931: gap of 100 bp
* 26932 27248: contig of 317 bp in length
* 27249 27348: gap of 100 bp
* 27349 28595: contig of 1247 bp in length
* 28596 28695: gap of 100 bp
* 28696 30616: contig of 1921 bp in length
* 30617 30716: gap of 100 bp
* 30717 32613: contig of 1897 bp in length
* 32614 32713: gap of 100 bp
* 32714 33958: contig of 1245 bp in length
* 33959 34058: gap of 100 bp
* 34059 35732: contig of 1674 bp in length
* 35733 35832: gap of 100 bp
* 35833 37422: contig of 1590 bp in length
* 37423 37522: gap of 100 bp
* 37523 38306: contig of 784 bp in length
* 38307 38406: gap of 100 bp
* 38407 40880: contig of 2474 bp in length
* 40881 40980: gap of 100 bp
* 40981 42955: contig of 1975 bp in length
* 42956 43055: gap of 100 bp
* 43056 45418: contig of 2363 bp in length
* 45419 45518: gap of 100 bp
* 45519 48193: contig of 2675 bp in length
* 48194 48293: gap of 100 bp
* 48294 50938: contig of 2645 bp in length
* 50939 51038: gap of 100 bp
* 51039 54450: contig of 3412 bp in length
* 54451 54550: gap of 100 bp
* 54551 57303: contig of 2753 bp in length
* 57304 57403: gap of 100 bp
* 57404 60335: contig of 2932 bp in length
* 60336 60435: gap of 100 bp
* 60436 64090: contig of 3655 bp in length
* 64091 64190: gap of 100 bp
* 64191 67400: contig of 3210 bp in length
* 67401 67500: gap of 100 bp
* 67501 70359: contig of 2859 bp in length
* 70360 70459: gap of 100 bp
* 70460 74596: contig of 4137 bp in length
* 74597 74696: gap of 100 bp
* 74697 78944: contig of 4248 bp in length
* 78945 79044: gap of 100 bp
* 79045 83836: contig of 4792 bp in length
* 83837 83936: gap of 100 bp
* 83937 89561: contig of 5625 bp in length
* 89562 89661: gap of 100 bp
* 89662 95570: contig of 5909 bp in length
* 95571 95670: gap of 100 bp
* 95671 100625: contig of 4955 bp in length


```

* 100626 100725: gap of 100 bp
* 100726 106496: contig of 5771 bp in length
* 106497 106596: gap of 100 bp
* 106597 113908: contig of 7312 bp in length
* 113909 114008: gap of 100 bp
* 114009 119847: contig of 5839 bp in length
* 119848 119947: gap of 100 bp
* 119948 126028: contig of 6081 bp in length
* 126029 126128: gap of 100 bp
* 126129 132814: contig of 6686 bp in length
* 132815 132914: gap of 100 bp
* 132915 139664: contig of 6750 bp in length
* 139665 139764: gap of 100 bp
* 139765 149495: contig of 9731 bp in length
* 149496 149595: gap of 100 bp
* 149596 157946: contig of 8353 bp in length
* 157949 158048: gap of 100 bp
* 158049 167045: contig of 8997 bp in length
* 167046 167145: gap of 100 bp
* 167146 181853: contig of 14708 bp in length.

```

FEATURES

source

```

1. .181853
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1"
/clone="RP11-153f19"
/clone_id="RPC1-11 Human Male BAC"
1. .1225
/misc_feature
/note="assembly_fragment"
1326. .2383
/misc_feature
/note="assembly_fragment"
2484. .3638
/misc_feature
/note="assembly_fragment"
3739. .4871
/misc_feature
/note="assembly_fragment"
4972. .6453
/misc_feature
/note="assembly_fragment"
6554. .7651
/misc_feature
/note="assembly_fragment"
7752. .8917
/misc_feature
/note="assembly_fragment"
9018. .10865
/misc_feature
/note="assembly_fragment"
10966. .12050
/misc_feature
/note="assembly_fragment"
12151. .13291
/misc_feature
/note="assembly_fragment"
13392. .14491
/misc_feature
/note="assembly_fragment"
14592. .15649

```

Query Match 1.9%; Score 19; DB 68; Length 181853;

Best Local Similarity 100.0%; Pred. No. 32;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 484 ttgttgagctaatagtg 502

Db 142035 TTGGTGGACGCTAATAGTG 142053

Search completed: September 21, 2001, 21:50:28
Job time: 28248 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 16:24:34 ; Search time 646.95 Seconds
(without alignments)
947.263 Million cell updates/sec

Title: US-09-138-735-1_COPY_1232_2207

Perfect score: 976

Sequence: 1 caggtacagcgtacgagctt.....tgtgaacgagcgccattatgt 976

Scoring table:

OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 730101 seqs, 313950809 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

N_Geneseq_0601:*

1: /cgnl_9/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /cgnl_9/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /cgnl_9/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /cgnl_9/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /cgnl_9/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /cgnl_9/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /cgnl_9/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /cgnl_9/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /cgnl_9/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /cgnl_9/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /cgnl_9/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /cgnl_9/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /cgnl_9/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /cgnl_9/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /cgnl_9/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /cgnl_9/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /cgnl_9/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /cgnl_9/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /cgnl_9/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /cgnl_9/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /cgnl_9/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Length | ID | Description |
|------------|-------|--------------|-------|-------------|
| 1 | 976 | 100.0 | 3402 | 20 AAX84092 |
| 2 | 874 | 89.5 | 3402 | 17 AAT7310 |
| 3 | 22 | 2.3 | 22 | 17 AAT7313 |
| 4 | 22 | 2.3 | 22 | 20 AAX84095 |
| 5 | 21 | 2.2 | 21 | 17 AAT7314 |
| 6 | 21 | 2.2 | 21 | 17 AAT7315 |
| 7 | 21 | 2.2 | 21 | 20 AAX84096 |
| 8 | 21 | 2.2 | 21 | 20 AAX84097 |
| 9 | 19 | 1.9 | 11878 | 19 AAV30199 |
| 10 | 19 | 1.9 | 11883 | 19 AAV30198 |
| 11 | 18 | 1.8 | 18 | 17 AAT7312 |

| | | | | | | |
|---|----|----|-----|--------|----|----------|
| C | 12 | 18 | 1.8 | 18 | 20 | AAX84094 |
| C | 13 | 18 | 1.8 | 454 | 21 | AAF09309 |
| C | 14 | 18 | 1.8 | 479 | 21 | AAC04783 |
| C | 15 | 18 | 1.8 | 486 | 21 | AAC36458 |
| C | 16 | 18 | 1.8 | 634 | 21 | AAC31869 |
| C | 17 | 18 | 1.8 | 866 | 20 | AAX39943 |
| C | 18 | 18 | 1.8 | 970 | 21 | AAC41168 |
| C | 19 | 18 | 1.8 | 1302 | 21 | AAC59435 |
| C | 20 | 18 | 1.8 | 2520 | 15 | AA077884 |
| C | 21 | 18 | 1.8 | 2520 | 17 | AAT27766 |
| C | 22 | 18 | 1.8 | 68750 | 21 | AAZ35887 |
| C | 23 | 18 | 1.8 | 71989 | 21 | AAA29349 |
| C | 24 | 18 | 1.7 | 198 | 20 | AAX84320 |
| C | 25 | 17 | 1.7 | 334 | 13 | AAQ26088 |
| C | 26 | 17 | 1.7 | 353 | 13 | AAQ09227 |
| C | 27 | 17 | 1.7 | 364 | 21 | AAC00184 |
| C | 28 | 17 | 1.7 | 386 | 21 | AAC01025 |
| C | 29 | 17 | 1.7 | 513 | 21 | AAC38084 |
| C | 30 | 17 | 1.7 | 552 | 21 | AAC93440 |
| C | 31 | 17 | 1.7 | 662 | 21 | AAE13645 |
| C | 32 | 17 | 1.7 | 673 | 19 | AAV16884 |
| C | 33 | 17 | 1.7 | 673 | 20 | AAZ26019 |
| C | 34 | 17 | 1.7 | 673 | 20 | AAZ87504 |
| C | 35 | 17 | 1.7 | 697 | 20 | AAZ33534 |
| C | 36 | 17 | 1.7 | 763 | 21 | AAC39218 |
| C | 37 | 17 | 1.7 | 800 | 21 | AAC99839 |
| C | 38 | 17 | 1.7 | 1148 | 21 | AAC42507 |
| C | 39 | 17 | 1.7 | 1273 | 21 | AAC49984 |
| C | 40 | 17 | 1.7 | 1277 | 21 | AAC37359 |
| C | 41 | 17 | 1.7 | 1338 | 19 | AAV64526 |
| C | 42 | 17 | 1.7 | 1338 | 20 | AAV44417 |
| C | 43 | 17 | 1.7 | 1338 | 19 | AAZ19327 |
| C | 44 | 17 | 1.7 | 1338 | 20 | AAZ19115 |
| C | 45 | 17 | 1.7 | 1470 | 21 | AAC78035 |
| C | 46 | 17 | 1.7 | 1647 | 21 | AAC79935 |
| C | 47 | 17 | 1.7 | 1949 | 21 | AAA26361 |
| C | 48 | 17 | 1.7 | 2457 | 22 | AAA91185 |
| C | 49 | 17 | 1.7 | 2545 | 20 | AAI14998 |
| C | 50 | 17 | 1.7 | 2545 | 21 | AAI74878 |
| C | 51 | 17 | 1.7 | 3337 | 17 | AAT34620 |
| C | 52 | 17 | 1.7 | 3337 | 20 | AAI15174 |
| C | 53 | 17 | 1.7 | 6171 | 21 | AAC74895 |
| C | 54 | 17 | 1.7 | 11212 | 21 | AAA58762 |
| C | 55 | 17 | 1.7 | 11672 | 21 | AAD01009 |
| C | 56 | 17 | 1.7 | 15611 | 22 | AAD01008 |
| C | 57 | 17 | 1.7 | 240825 | 12 | AAF24497 |
| C | 58 | 16 | 1.6 | 31 | 14 | AAQ26748 |
| C | 59 | 16 | 1.6 | 41 | 15 | AAQ36842 |
| C | 60 | 16 | 1.6 | 222 | 14 | AAQ73710 |
| C | 61 | 16 | 1.6 | 334 | 20 | AAZ25170 |
| C | 62 | 16 | 1.6 | 369 | 21 | AAC03019 |
| C | 63 | 16 | 1.6 | 371 | 20 | AAZ21399 |
| C | 64 | 16 | 1.6 | 398 | 21 | AAC01286 |
| C | 65 | 16 | 1.6 | 413 | 21 | AAV79394 |
| C | 66 | 16 | 1.6 | 449 | 21 | AAC01726 |
| C | 67 | 16 | 1.6 | 541 | 19 | AAT76907 |
| C | 68 | 16 | 1.6 | 598 | 21 | AAA50408 |
| C | 69 | 16 | 1.6 | 624 | 18 | AAT84039 |
| C | 70 | 16 | 1.6 | 651 | 19 | AAV15445 |
| C | 71 | 16 | 1.6 | 717 | 21 | AAA61659 |
| C | 72 | 16 | 1.6 | 720 | 22 | AAV71218 |
| C | 73 | 16 | 1.6 | 774 | 19 | AAV28667 |
| C | 74 | 16 | 1.6 | 774 | 20 | AAZ15622 |
| C | 75 | 16 | 1.6 | 865 | 21 | AAZ98109 |
| C | 76 | 16 | 1.6 | 898 | 21 | AAC99832 |
| C | 77 | 16 | 1.6 | 928 | 20 | AAX61395 |
| C | 78 | 16 | 1.6 | 1046 | 21 | AAC36246 |
| C | 79 | 16 | 1.6 | 1049 | 21 | AAC49065 |
| C | 80 | 16 | 1.6 | 1185 | 19 | AAV05039 |
| C | 81 | 16 | 1.6 | 1185 | 20 | AAZ08784 |
| C | 82 | 16 | 1.6 | 1185 | 21 | AAV1692 |
| C | 83 | 16 | 1.6 | 1234 | 21 | AAV53983 |
| C | 84 | 16 | 1.6 | 1314 | 19 | AAV00732 |

PCR primer for T.
Fusarium venenatum
Human secreted pro
Arabidopsis thalia
Arabidopsis thalia
Gastric cancer ass
Arabidopsis thalia
Human secreted pro
Neural thread prot
AD 16c human neutra
Sorantium cellulos
Stealth virus nucl
MIG-2. Mammalian.
Human secreted pro
Human secreted pro
Human secreted pro
Arabidopsis thalia
Human secreted pro
Aspergillus oryzae
Human prostate can
Prostate disease m
Prostate, breast a
Human prostate can
Arabidopsis thalia
Human secreted pro
Arabidopsis thalia
Arabidopsis thalia
M. tuberculosis im
Mycobacterium tube
M. tuberculosis an
M. tuberculosis re
Human cancer assoc
Human secreted pro
Human secreted pro
zea mays ZmEIN3-2
DNA encoding a mon
Human chemokine co
P. vitax ESP-1 blo
DNA encoding a sec
Human ORFX ORF450
DNA encoding a cyc
Escherichia coli p
Escherichia coli p
Human PG-3 gene.
Protease-A gene pr
Adaptor-primer. S
Retrotransposon pro
HIV-1 group O isol
Human secreted pro
3' Fragment of Abs
Human secreted pro
Eucalyptus grandis
Human secreted pro
S. glaucusens Hst
Human increased in
DNA encoding a Sta
Human gene fragmen
cDNA encoding mous
Corynebacterium gl
Ripening banana pu
Human gene express
Human signal pepti
Human secreted pro
DNA encoding a hum
Arabidopsis thalia
Arabidopsis thalia
Alcaligenes eutrop
A. eutrophus beta-
A. eutrophus bktB
Arabidopsis thalia
Aspergillus terreu

| | | | | | | | | | | | | | |
|-----|----|-----|---------|----|-----------|---------------------|-------|----|-----|-----|----|----------|---------------------|
| 85 | 16 | 1.6 | 1314 | 19 | AAV09626 | A. terreus FAOD-L | c 158 | 15 | 1.5 | 44 | 18 | AAV58825 | Glyceryl DNP/DNP D |
| 86 | 16 | 1.6 | 1314 | 20 | AAZ07721 | Heat-resistant fru | c 159 | 15 | 1.5 | 81 | 16 | AAV06202 | HIV-1 reverse tran |
| 87 | 16 | 1.6 | 1411 | 14 | AAQ41260 | Encodes repressor | c 160 | 15 | 1.5 | 157 | 18 | AAV84365 | Friedreich's ataxi |
| 88 | 16 | 1.6 | 1411 | 15 | AAQ72717 | Dialkylglycine dec | c 161 | 15 | 1.5 | 168 | 21 | AAA44666 | Human secreted exp |
| 89 | 16 | 1.6 | 1464 | 22 | AAV71799 | Corynebacterium g1 | c 162 | 15 | 1.5 | 196 | 21 | AAA41667 | Human secreted exp |
| 90 | 16 | 1.6 | 1560 | 20 | AAZ27334 | Human secreted pro | c 163 | 15 | 1.5 | 226 | 19 | AAV15050 | Xylanase activity |
| 91 | 16 | 1.6 | 1569 | 20 | AAA25204 | Maize cinnamate-4- | c 164 | 15 | 1.5 | 295 | 16 | AAV21139 | Human gene signalu |
| 92 | 16 | 1.6 | 1685 | 21 | AAA61660 | CDNA encoding mous | c 165 | 15 | 1.5 | 297 | 21 | AAV77752 | CDNA encoding huma |
| 93 | 16 | 1.6 | 1722 | 16 | AAV00402 | Squalene epoxidase | c 166 | 15 | 1.5 | 300 | 20 | AAV14302 | Human gene express |
| 94 | 16 | 1.6 | 1734 | 17 | AAV45893 | R. capsulatus acet | c 167 | 15 | 1.5 | 300 | 20 | AAV12540 | Human gene express |
| 95 | 16 | 1.6 | 1810 | 22 | AAA54510 | Fructan exohydrola | c 168 | 15 | 1.5 | 300 | 21 | AAV00491 | Human colon cancer |
| 96 | 16 | 1.6 | 1816 | 22 | AAV72758 | Human prostate can | c 169 | 15 | 1.5 | 304 | 21 | AAV28635 | Human secreted pro |
| 97 | 16 | 1.6 | 1864 | 20 | AAZ24942 | Wild type A. eulio | c 170 | 15 | 1.5 | 325 | 20 | AAV86047 | Human secreted pro |
| 98 | 16 | 1.6 | 2040 | 21 | AAZ46364 | Haemorrhagic enter | c 171 | 15 | 1.5 | 325 | 21 | AAV60344 | Exon 1 of the huma |
| 99 | 16 | 1.6 | 2063 | 22 | AAV84244 | Signal transductio | c 172 | 15 | 1.5 | 325 | 21 | AAV66670 | Human neuropeptide |
| 100 | 16 | 1.6 | 2068 | 21 | AAA61661 | CDNA encoding mous | c 173 | 15 | 1.5 | 328 | 21 | AAV12718 | Human secreted pro |
| 101 | 16 | 1.6 | 2070 | 21 | AAA61662 | CDNA encoding mous | c 174 | 15 | 1.5 | 328 | 22 | AAV93718 | CDNA encoding SRT |
| 102 | 16 | 1.6 | 2190 | 14 | AAQ37237 | gdhA gene of C.mel | c 175 | 15 | 1.5 | 384 | 22 | AAV65140 | Novel human polynu |
| 103 | 16 | 1.6 | 2214 | 19 | AAV22682 | New DNA sequence i | c 176 | 15 | 1.5 | 395 | 22 | AAV66688 | Novel human polynu |
| 104 | 16 | 1.6 | 2274 | 21 | AAV47059 | Zea mays DNA fragm | c 177 | 15 | 1.5 | 397 | 19 | AAV44835 | Clone HMSCX46R rel |
| 105 | 16 | 1.6 | 2310 | 20 | AAZ21502 | Saccharopolyspora | c 178 | 15 | 1.5 | 397 | 22 | AAV65139 | Novel human polynu |
| 106 | 16 | 1.6 | 2499 | 21 | AAV64076 | Mouse ion channel | c 179 | 15 | 1.5 | 399 | 22 | AAV64930 | Novel human polynu |
| 107 | 16 | 1.6 | 2535 | 21 | AAA55201 | C. symbiosum open | c 180 | 15 | 1.5 | 403 | 22 | AAV66880 | Novel human polynu |
| 108 | 16 | 1.6 | 2610 | 21 | AAV290578 | Human death induce | c 181 | 15 | 1.5 | 419 | 21 | AAA43539 | Novel human polynu |
| 109 | 16 | 1.6 | 2886 | 21 | AAZ51244 | Murine TANGO 130 c | c 182 | 15 | 1.5 | 426 | 20 | AAV08739 | Nucleotide encodin |
| 110 | 16 | 1.6 | 2891 | 22 | AAV71640 | Corynebacterium g1 | c 183 | 15 | 1.5 | 430 | 21 | AAA43600 | Human secreted exp |
| 111 | 16 | 1.6 | 3045 | 20 | AAZ12185 | Neisseria meningit | c 184 | 15 | 1.5 | 462 | 20 | AAV20850 | Polynucleotide seq |
| 112 | 16 | 1.6 | 3059 | 21 | AAV60191 | Neisseria gonorrhoe | c 185 | 15 | 1.5 | 469 | 19 | AAV59518 | Human secreted pro |
| 113 | 16 | 1.6 | 3059 | 21 | AAV60191 | Hydrophobic domain | c 186 | 15 | 1.5 | 483 | 21 | AAV79184 | Human secreted pro |
| 114 | 16 | 1.6 | 3148 | 21 | AAV67639 | Human ORFX ORF1794 | c 187 | 15 | 1.5 | 499 | 21 | AAV09046 | Human lung tumour- |
| 115 | 16 | 1.6 | 3153 | 19 | AAV19895 | Glucodextranase ge | c 188 | 15 | 1.5 | 500 | 21 | AAV65460 | Fusarium venenatum |
| 116 | 16 | 1.6 | 3240 | 21 | AAZ98339 | A. thaliana gene 1 | c 189 | 15 | 1.5 | 512 | 19 | AAV29356 | Porcine BAC-PRGF2- |
| 117 | 16 | 1.6 | 3331 | 19 | AAZ22683 | New DNA sequence i | c 190 | 15 | 1.5 | 574 | 21 | AAV56056 | Calcium ion channe |
| 118 | 16 | 1.6 | 3331 | 20 | AAZ32021 | Human METL1 relate | c 191 | 15 | 1.5 | 603 | 22 | AAV22775 | Eucaulyptus grandis |
| 119 | 16 | 1.6 | 3331 | 22 | AAV098078 | D86074 CDNA clone. | c 192 | 15 | 1.5 | 628 | 22 | AAV22779 | Human prostate can |
| 120 | 16 | 1.6 | 4394 | 17 | AAV098098 | Alpha-ketoglutaric | c 193 | 15 | 1.5 | 631 | 21 | AAV54906 | Human prostate can |
| 121 | 16 | 1.6 | 4384 | 19 | AAV13157 | B. lactofementum | c 194 | 15 | 1.5 | 633 | 21 | AAV36696 | Arabidopsis thalia |
| 122 | 16 | 1.6 | 4867 | 16 | AAV39170 | Erylnia carotovora | c 195 | 15 | 1.5 | 633 | 21 | AAV31909 | Arabidopsis thalia |
| 123 | 16 | 1.6 | 5117 | 16 | AAV06027 | hubeP-1 CDNA. Hom | c 196 | 15 | 1.5 | 637 | 21 | AAV37714 | Arabidopsis thalia |
| 124 | 16 | 1.6 | 6242 | 21 | AAV94945 | Human CDNA encodin | c 197 | 15 | 1.5 | 653 | 21 | AAV53412 | Neisseria meningit |
| 125 | 16 | 1.6 | 6545 | 17 | AAV27537 | Maltese binding pr | c 198 | 15 | 1.5 | 657 | 21 | AAV13288 | Aspergillus oryzae |
| 126 | 16 | 1.6 | 7732 | 20 | AAV77723 | PDVZ vector. Synt | c 199 | 15 | 1.5 | 657 | 21 | AAV53410 | Neisseria gonorrhoe |
| 127 | 16 | 1.6 | 9243 | 21 | AAV61534 | N. meningitidis pa | c 200 | 15 | 1.5 | 658 | 21 | AAV54905 | Arabidopsis thalia |
| 128 | 16 | 1.6 | 9551 | 20 | AAV22301 | CDNA encoding a hu | c 201 | 15 | 1.5 | 659 | 21 | AAV81591 | N. meningitidis pa |
| 129 | 16 | 1.6 | 10391 | 20 | AAV13070 | Enterococcus faeca | c 202 | 15 | 1.5 | 672 | 21 | AAV52269 | Arabidopsis thalia |
| 130 | 16 | 1.6 | 16310 | 21 | AAV21086 | Human low adenosin | c 203 | 15 | 1.5 | 674 | 21 | AAV02040 | Human colon cancer |
| 131 | 16 | 1.6 | 16310 | 21 | AAV34964 | Human adenosine re | c 204 | 15 | 1.5 | 675 | 22 | AAV22781 | Human prostate can |
| 132 | 16 | 1.6 | 17634 | 21 | AAV21087 | Human low adenosin | c 205 | 15 | 1.5 | 679 | 21 | AAV13170 | Aspergillus oryzae |
| 133 | 16 | 1.6 | 17634 | 21 | AAV34965 | Human adenosine re | c 206 | 15 | 1.5 | 686 | 21 | AAV60054 | Human secreted pro |
| 134 | 16 | 1.6 | 17920 | 21 | AAV39168 | Erylnia carotovora | c 207 | 15 | 1.5 | 689 | 21 | AAV66714 | Polynucleotide iso |
| 135 | 16 | 1.6 | 21721 | 20 | AAV83427 | Human lipolysis st | c 208 | 15 | 1.5 | 699 | 21 | AAV13007 | Aspergillus oryzae |
| 136 | 16 | 1.6 | 22976 | 20 | AAV83426 | Genomic region con | c 209 | 15 | 1.5 | 752 | 21 | AAV77117 | Human ORFX ORF2672 |
| 137 | 16 | 1.6 | 23187 | 21 | AAV50273 | Human lipolysis st | c 210 | 15 | 1.5 | 791 | 21 | AAV54902 | Arabidopsis thalia |
| 138 | 16 | 1.6 | 23187 | 22 | AAV62331 | Human leptin fragm | c 211 | 15 | 1.5 | 792 | 21 | AAV52503 | Arabidopsis thalia |
| 139 | 16 | 1.6 | 26270 | 21 | AAV46355 | Haemorrhagic enter | c 212 | 15 | 1.5 | 799 | 20 | AAV16224 | Human gene express |
| 140 | 16 | 1.6 | 35133 | 21 | AAV81518 | N. meningitidis pa | c 213 | 15 | 1.5 | 800 | 19 | AAV63926 | Mycobacterium tube |
| 141 | 16 | 1.6 | 39796 | 21 | AAV61681 | Nucleotide sequenc | c 214 | 15 | 1.5 | 800 | 20 | AAV81033 | Nucleotide sequenc |
| 142 | 16 | 1.6 | 42432 | 21 | AAV55187 | Cenarchaeum symbio | c 215 | 15 | 1.5 | 810 | 19 | AAV00257 | Nucleotide sequenc |
| 143 | 16 | 1.6 | 47475 | 21 | AAV81465 | N. meningitidis pa | c 216 | 15 | 1.5 | 810 | 21 | AAV39970 | Murine TANGO 188 c |
| 144 | 16 | 1.6 | 72928 | 20 | AAV18355 | Human AS7H1J 5' ge | c 217 | 15 | 1.5 | 819 | 15 | AAV58817 | NANBH virus gene f |
| 145 | 16 | 1.6 | 72928 | 21 | AAV80253 | Human AS7H1J 5' ge | c 218 | 15 | 1.5 | 833 | 20 | AAV27324 | Human secreted pro |
| 146 | 16 | 1.6 | 102634 | 21 | AAV81464 | N. meningitidis pa | c 219 | 15 | 1.5 | 837 | 21 | AAV30334 | Nucleotide sequenc |
| 147 | 16 | 1.6 | 349980 | 21 | AAV21608 | Neisseria meningit | c 220 | 15 | 1.5 | 838 | 16 | AAV09075 | Human beta-kinesin |
| 148 | 16 | 1.6 | 349980 | 21 | AAV21609 | Neisseria meningit | c 221 | 15 | 1.5 | 841 | 22 | AAV75381 | Ralstonia eutropha |
| 149 | 16 | 1.6 | 349980 | 21 | AAV21610 | Neisseria meningit | c 222 | 15 | 1.5 | 846 | 20 | AAV25129 | Soybean chalcone 1 |
| 150 | 16 | 1.6 | 349980 | 21 | AAV21611 | Neisseria meningit | c 223 | 15 | 1.5 | 846 | 21 | AAV38424 | Sequence of region |
| 151 | 16 | 1.6 | 349980 | 21 | AAV21612 | Neisseria meningit | c 224 | 15 | 1.5 | 876 | 8 | AAV70094 | DNA encoding a hum |
| 152 | 16 | 1.6 | 837096 | 21 | AAV81489 | N. meningitidis pa | c 225 | 15 | 1.5 | 901 | 20 | AAV86270 | DNA encoding a hum |
| 153 | 16 | 1.6 | 837096 | 21 | AAV81489 | N. meningitidis pa | c 226 | 15 | 1.5 | 909 | 21 | AAV58817 | Arabidopsis thalia |
| 154 | 16 | 1.6 | 1437066 | 21 | AAV81490 | N. meningitidis B | c 227 | 15 | 1.5 | 921 | 21 | AAV45812 | DNA encoding a hum |
| 155 | 15 | 1.5 | 40 | 18 | AAV14833 | PCR primer HG52.16 | c 228 | 15 | 1.5 | 943 | 21 | AAV58815 | Clone HYASC03 codi |
| 156 | 15 | 1.5 | 40 | 18 | AAV58818 | DNP Du labelled ol | c 229 | 15 | 1.5 | 948 | 22 | AAV90012 | Murine Mantic Frlng |
| 157 | 15 | 1.5 | 44 | 18 | AAV58819 | Glyceryl-based DNP | c 230 | 15 | 1.5 | 966 | 19 | AAV36218 | |

| | | | | | | | | | | | | | |
|-----|----|-----|------|----|----------|--------------------|-----|----|-----|------|----|----------|---------------------|
| 231 | 15 | 1.5 | 1005 | 21 | AAc51280 | Arabidopsis thalia | 304 | 15 | 1.5 | 1608 | 14 | AAQ41461 | Mycobacterium para |
| 232 | 15 | 1.5 | 1014 | 21 | AAc34499 | Arabidopsis thalia | 305 | 15 | 1.5 | 1631 | 21 | AAZ98141 | Human signal pepti |
| 233 | 15 | 1.5 | 1029 | 17 | AAr72218 | G protein coupled | 306 | 15 | 1.5 | 1632 | 14 | AAQ42976 | Carnitine O-m |
| 234 | 15 | 1.5 | 1065 | 21 | AAr14557 | Aspergillus oryzae | 307 | 15 | 1.5 | 1636 | 18 | AAr85271 | HMG-CoA reductase |
| 235 | 15 | 1.5 | 1068 | 21 | AAc48553 | Arabidopsis thalia | 308 | 15 | 1.5 | 1636 | 19 | AAV49559 | Human liver cell c |
| 236 | 15 | 1.5 | 1078 | 20 | AAx08745 | Nucleotide encodin | 309 | 15 | 1.5 | 1667 | 21 | AAZ51207 | Rat lost in leukae |
| 237 | 15 | 1.5 | 1102 | 22 | AAr22774 | Human prostate can | 310 | 15 | 1.5 | 1669 | 20 | AAZ42020 | Human endometrium |
| 238 | 15 | 1.5 | 1129 | 21 | AAc54524 | Arabidopsis thalia | 311 | 15 | 1.5 | 1704 | 15 | AAV44509 | Human TNF receptor |
| 239 | 15 | 1.5 | 1144 | 21 | AAc33490 | Arabidopsis thalia | 312 | 15 | 1.5 | 1704 | 20 | AAV83763 | Human tumour necro |
| 240 | 15 | 1.5 | 1152 | 21 | AAc43754 | Zea mays DNA fragm | 313 | 15 | 1.5 | 1704 | 21 | AAZ8135 | Human TR2 receptor |
| 241 | 15 | 1.5 | 1159 | 21 | AAc77857 | Human cancer assoc | 314 | 15 | 1.5 | 1707 | 22 | AAr72809 | Secreted protein g |
| 242 | 15 | 1.5 | 1189 | 21 | AAr07514 | Fusarium venenatum | 315 | 15 | 1.5 | 1715 | 16 | AAr04588 | Rat Vh-1 CDNA. R |
| 243 | 15 | 1.5 | 1200 | 22 | AAr71690 | Corynebacterium g1 | 316 | 15 | 1.5 | 1738 | 11 | AAZ51206 | Rat lost in leukae |
| 244 | 15 | 1.5 | 1208 | 21 | AAc48562 | Arabidopsis thalia | 317 | 15 | 1.5 | 1739 | 21 | AAc53432 | Arabidopsis thalia |
| 245 | 15 | 1.5 | 1211 | 21 | AAc41977 | Arabidopsis thalia | 318 | 15 | 1.5 | 1765 | 20 | AAZ06232 | Human secreted pro |
| 246 | 15 | 1.5 | 1213 | 21 | AAc40267 | Arabidopsis thalia | 319 | 15 | 1.5 | 1777 | 16 | AAQ93368 | S. lividans protea |
| 247 | 15 | 1.5 | 1221 | 21 | AAr97139 | Human gene 75 DNA | 320 | 15 | 1.5 | 1777 | 20 | AAV84067 | Clone PB-2 encodin |
| 248 | 15 | 1.5 | 1223 | 21 | AAc99054 | Human pancreatic c | 321 | 15 | 1.5 | 1779 | 21 | AAc61405 | CDNA encoding a su |
| 249 | 15 | 1.5 | 1242 | 21 | AAc53434 | Arabidopsis thalia | 322 | 15 | 1.5 | 1779 | 20 | AAZ12073 | Neisseria meningit |
| 250 | 15 | 1.5 | 1249 | 19 | AAV38987 | CDNA encoding a hu | 323 | 15 | 1.5 | 1779 | 20 | AAZ12074 | Neisseria meningit |
| 251 | 15 | 1.5 | 1249 | 19 | AAV38987 | CDNA encoding a hu | 324 | 15 | 1.5 | 1779 | 20 | AAZ12075 | Neisseria gonorrh |
| 252 | 15 | 1.5 | 1263 | 21 | AAZ35255 | Plant retroelement | 325 | 15 | 1.5 | 1779 | 20 | AAZ12075 | Neisseria gonorrh |
| 253 | 15 | 1.5 | 1265 | 21 | AAc42091 | Arabidopsis thalia | 326 | 15 | 1.5 | 1779 | 21 | AAZ53673 | Neisseria meningit |
| 254 | 15 | 1.5 | 1268 | 21 | AAZ51410 | A. thaliana VB33 c | 327 | 15 | 1.5 | 1779 | 21 | AAZ53675 | Neisseria meningit |
| 255 | 15 | 1.5 | 1275 | 21 | AAc60190 | Streptococcus pneu | 328 | 15 | 1.5 | 1780 | 21 | AAZ65335 | Human secreted pro |
| 256 | 15 | 1.5 | 1275 | 22 | AAr27608 | Mevlonate pathway | 329 | 15 | 1.5 | 1788 | 22 | AAr32651 | Human CDNA encodin |
| 257 | 15 | 1.5 | 1278 | 20 | AAr90824 | DNA encoding human | 330 | 15 | 1.5 | 1827 | 21 | AAc59682 | Novel Human protei |
| 258 | 15 | 1.5 | 1281 | 20 | AAK15146 | Coding region for | 331 | 15 | 1.5 | 1831 | 20 | AAZ6880 | DNA encoding a pro |
| 259 | 15 | 1.5 | 1287 | 22 | AAr67832 | Corynebacterium g1 | 332 | 15 | 1.5 | 1831 | 21 | AAZ09889 | Human OCTN2 CDNA s |
| 260 | 15 | 1.5 | 1290 | 19 | AAV59676 | Human secreted pro | 333 | 15 | 1.5 | 1855 | 22 | AAr30190 | Clone 11696905-0-4 |
| 261 | 15 | 1.5 | 1294 | 20 | AAZ27438 | Human secreted pro | 334 | 15 | 1.5 | 1858 | 20 | AAZ55599 | DNA encoding a hum |
| 262 | 15 | 1.5 | 1294 | 21 | AAr13611 | Aspergillus oryzae | 335 | 15 | 1.5 | 1862 | 22 | AAZ02606 | Human angiotensin |
| 263 | 15 | 1.5 | 1313 | 16 | AAQ91637 | Mouse sonic hedgeh | 336 | 15 | 1.5 | 1863 | 21 | AAZ8268 | Yeast essential ge |
| 264 | 15 | 1.5 | 1313 | 20 | AAZ25620 | Mouse sonic hedgeh | 337 | 15 | 1.5 | 1869 | 20 | AAZ33970 | Human TIE ligand N |
| 265 | 15 | 1.5 | 1313 | 20 | AAZ25101 | Mouse sonic hedgeh | 338 | 15 | 1.5 | 1869 | 21 | AAZ60537 | Human TIE ligand N |
| 266 | 15 | 1.5 | 1313 | 20 | AAx07274 | Mouse sonic hedgeh | 339 | 15 | 1.5 | 1869 | 21 | AAZ50126 | Maize MLO3 CDNA. |
| 267 | 15 | 1.5 | 1313 | 20 | AAx16185 | Mouse Shh hedgehog | 340 | 15 | 1.5 | 1869 | 22 | AAc97388 | Human angiotensin |
| 268 | 15 | 1.5 | 1313 | 21 | AAZ50449 | Mouse sonic hedgeh | 341 | 15 | 1.5 | 1873 | 22 | AAr27220 | Human neovasculari |
| 269 | 15 | 1.5 | 1313 | 21 | AAZ27879 | Mouse sonic hedgeh | 342 | 15 | 1.5 | 1876 | 12 | AAQ14498 | Clone 1 for trunca |
| 270 | 15 | 1.5 | 1313 | 21 | AAZ30277 | Partial mouse Shh | 343 | 15 | 1.5 | 1876 | 12 | AAQ14499 | Clone 2 for trunca |
| 271 | 15 | 1.5 | 1313 | 21 | AAZ52260 | Mouse sonic hedgeh | 344 | 15 | 1.5 | 1876 | 20 | AAZ15144 | CDNA encoding huma |
| 272 | 15 | 1.5 | 1313 | 22 | AAc87077 | Nucleotide sequenc | 345 | 15 | 1.5 | 1879 | 21 | AAZ45850 | Human muscle angio |
| 273 | 15 | 1.5 | 1313 | 22 | AAr27016 | Mouse sonic hedgeh | 346 | 15 | 1.5 | 1879 | 22 | AAr29701 | Human angiotensin |
| 274 | 15 | 1.5 | 1314 | 22 | AAZ37897 | Mouse sonic hedgeh | 347 | 15 | 1.5 | 1888 | 19 | AAV49558 | Human liver cell c |
| 275 | 15 | 1.5 | 1330 | 21 | AAZ54055 | Arabidopsis thalia | 348 | 15 | 1.5 | 1894 | 20 | AAZ87858 | Human fibrinogen d |
| 276 | 15 | 1.5 | 1340 | 21 | AAZ26964 | Human coenzyme A-u | 349 | 15 | 1.5 | 1899 | 21 | AAZ59302 | Insulin-like growt |
| 277 | 15 | 1.5 | 1359 | 19 | AAV41924 | Nucleotide sequenc | 350 | 15 | 1.5 | 1911 | 21 | AAr22306 | Human secreted pro |
| 278 | 15 | 1.5 | 1362 | 21 | AAZ50399 | Human angiotensin | 351 | 15 | 1.5 | 1913 | 22 | AAZ97138 | Human gene 75 DNA |
| 279 | 15 | 1.5 | 1362 | 21 | AAZ35278 | Soybean retroelme | 352 | 15 | 1.5 | 1921 | 22 | AAZ44835 | RPP5-like protein |
| 280 | 15 | 1.5 | 1368 | 21 | AAZ54153 | Arabidopsis thalia | 353 | 15 | 1.5 | 1931 | 21 | AAZ97093 | Human secreted pro |
| 281 | 15 | 1.5 | 1368 | 21 | AAZ53670 | Neisseria gonorrhe | 354 | 15 | 1.5 | 1938 | 13 | AAQ22341 | Sequence encoding |
| 282 | 15 | 1.5 | 1368 | 21 | AAZ53672 | Neisseria meningit | 355 | 15 | 1.5 | 1938 | 13 | AAQ22341 | Sequence encoding |
| 283 | 15 | 1.5 | 1368 | 22 | AAr72256 | Corynebacterium g1 | 356 | 15 | 1.5 | 1947 | 21 | AAZ56176 | Human adipophillin- |
| 284 | 15 | 1.5 | 1368 | 22 | AAr72257 | Corynebacterium g1 | 357 | 15 | 1.5 | 1955 | 21 | AAZ29588 | CDNA encoding a hu |
| 285 | 15 | 1.5 | 1374 | 21 | AAc51258 | Arabidopsis thalia | 358 | 15 | 1.5 | 1955 | 21 | AAZ36987 | DNA encoding a hu |
| 286 | 15 | 1.5 | 1376 | 21 | AAc33903 | Arabidopsis thalia | 359 | 15 | 1.5 | 1968 | 21 | AAr18104 | Lung cancer associ |
| 287 | 15 | 1.5 | 1376 | 22 | AAc85554 | CDNA encoding CDIF | 360 | 15 | 1.5 | 1973 | 21 | AAc60998 | Human vesicle asso |
| 288 | 15 | 1.5 | 1379 | 21 | AAZ97132 | Human secreted pro | 361 | 15 | 1.5 | 1980 | 18 | AAr59454 | H-Delta-1 contig c |
| 289 | 15 | 1.5 | 1381 | 20 | AAV84482 | Human secreted pro | 362 | 15 | 1.5 | 2004 | 21 | AAZ54137 | Arabidopsis thalia |
| 290 | 15 | 1.5 | 1390 | 22 | AAr60982 | P. putida KT2440-a | 363 | 15 | 1.5 | 2008 | 20 | AAZ41299 | Human normal ovar |
| 291 | 15 | 1.5 | 1396 | 21 | AAc33740 | Arabidopsis thalia | 364 | 15 | 1.5 | 2044 | 17 | AAr13167 | Rat interleukin-1 |
| 292 | 15 | 1.5 | 1422 | 22 | AAr60986 | P. putida KT2440-a | 365 | 15 | 1.5 | 2044 | 18 | AAr84010 | DNA encoding a sta |
| 293 | 15 | 1.5 | 1433 | 17 | AAr39049 | CDNA encoding cell | 366 | 15 | 1.5 | 2051 | 20 | AAZ84059 | Heparin sulphate |
| 294 | 15 | 1.5 | 1437 | 20 | AAZ84973 | Human secreted pro | 367 | 15 | 1.5 | 2051 | 21 | AAZ28059 | Human HS65T1 CDNA. |
| 295 | 15 | 1.5 | 1448 | 20 | AAZ42009 | Human endometrium | 368 | 15 | 1.5 | 2084 | 21 | AAZ52492 | Human secreted pro |
| 296 | 15 | 1.5 | 1466 | 21 | AAZ38414 | Arabidopsis thalia | 369 | 15 | 1.5 | 2093 | 21 | AAZ43805 | Human adult brain |
| 297 | 15 | 1.5 | 1498 | 21 | AAZ45265 | Arabidopsis thalia | 370 | 15 | 1.5 | 2104 | 22 | AAZ39969 | Murine TANC0 188 c |
| 298 | 15 | 1.5 | 1504 | 17 | AAr36177 | Murine int6 gene a | 371 | 15 | 1.5 | 2104 | 22 | AAZ37120 | Nucleotide sequenc |
| 299 | 15 | 1.5 | 1512 | 21 | AAZ64513 | Nucleotide sequenc | 372 | 15 | 1.5 | 2160 | 20 | AAZ86635 | Nucleotide sequenc |
| 300 | 15 | 1.5 | 1572 | 20 | AAx02671 | T. versicolor lacc | 373 | 15 | 1.5 | 2185 | 22 | AAZ90575 | Human PRO861 CDNA. |
| 301 | 15 | 1.5 | 1572 | 20 | AAx02672 | T. versicolor lacc | 374 | 15 | 1.5 | 2185 | 22 | AAZ91472 | Human PRO861 CDNA. |
| 302 | 15 | 1.5 | 1591 | 16 | AAQ97586 | DNA encoding prote | 375 | 15 | 1.5 | 2185 | 22 | AAZ91578 | Human PRO861 CDNA. |
| 303 | 15 | 1.5 | 1592 | 21 | AAc77050 | Human ORFX ORF2605 | 376 | 15 | 1.5 | 2187 | 21 | AAZ29701 | Wild-type human c- |

| | | | | | | | | | | | | | |
|-------|----|-----|------|----|-----------|----------------------|-------|----|-----|--------|----|-----------|---------------------|
| C 377 | 15 | 1.5 | 2200 | 22 | AAF27713 | Human transport pr | C 450 | 15 | 1.5 | 3867 | 21 | AA293224 | Fatty acid desatur |
| C 378 | 15 | 1.5 | 2221 | 20 | AAZ00370 | Partial nucleotide | 451 | 15 | 1.5 | 3981 | 19 | AAV40381 | Strawberry structu |
| C 379 | 15 | 1.5 | 2249 | 21 | AA37055 | Human PRO1374 (UNQ | 452 | 15 | 1.5 | 4030 | 20 | AA231717 | Human patched-2 co |
| C 380 | 15 | 1.5 | 2249 | 22 | AAF54285 | DNA encoding prote | 453 | 15 | 1.5 | 4066 | 19 | AAV69921 | Expression vector |
| C 381 | 15 | 1.5 | 2263 | 20 | AAV80628 | Kinase injury asso | 454 | 15 | 1.5 | 4201 | 21 | AACT6084 | Human cancer assoc |
| C 382 | 15 | 1.5 | 2271 | 11 | AAO06844 | Amylase gene from | 455 | 15 | 1.5 | 4263 | 20 | AAV72672 | Hamster 3' DNA seq |
| C 383 | 15 | 1.5 | 2276 | 18 | AAI86161 | PHA depolymerase g | 456 | 15 | 1.5 | 4263 | 21 | AAAI1431 | cDNA encoding mur |
| C 384 | 15 | 1.5 | 2288 | 21 | AAZ97061 | Human secreted pro | 457 | 15 | 1.5 | 4263 | 21 | AAAI1379 | Mouse MT5-MMP matr |
| C 385 | 15 | 1.5 | 2291 | 9 | AAAB0309 | Entire amyLase gen | 458 | 15 | 1.5 | 4296 | 20 | AAZ32436 | Rat TAO2 kinase en |
| C 386 | 15 | 1.5 | 2311 | 21 | AAFI8225 | Lung cancer associ | 459 | 15 | 1.5 | 4751 | 20 | AAZ09496 | Human heart tissue |
| C 387 | 15 | 1.5 | 2315 | 21 | AAAC90027 | Clone HFKCD20 codi | 460 | 15 | 1.5 | 4879 | 18 | AAV74400 | Staphylococcus aur |
| C 388 | 15 | 1.5 | 2324 | 21 | AAAI5288 | cDNA encoding prot | 461 | 15 | 1.5 | 4880 | 14 | AAO36660 | Sequence of the po |
| C 389 | 15 | 1.5 | 2332 | 21 | AAV72398 | Human nucleic acid | 462 | 15 | 1.5 | 4984 | 19 | AAV18986 | Poly(hydroxyl)Kha |
| C 390 | 15 | 1.5 | 2339 | 22 | AAAF92112 | Human PRO1753 cDNA | 463 | 15 | 1.5 | 5162 | 19 | AAV69919 | Expression vector |
| C 391 | 15 | 1.5 | 2420 | 21 | AAZ46489 | PKA substrate, Csk | 464 | 15 | 1.5 | 5162 | 19 | AAV69922 | Mutagenic plasmid |
| C 392 | 15 | 1.5 | 2501 | 21 | AAAS2499 | 2.5 kb human MN ge | 465 | 15 | 1.5 | 5262 | 19 | AAV69927 | Promoter probe vec |
| C 393 | 15 | 1.5 | 2501 | 21 | AAAI6580 | MN genomic region | 466 | 15 | 1.5 | 5342 | 21 | AAAD00374 | Rat smooth muscle- |
| C 394 | 15 | 1.5 | 2507 | 21 | AAZ53365 | NSBQ gene-17 assoc | 467 | 15 | 1.5 | 5375 | 22 | AAAC62160 | Nucleotide sequenc |
| C 395 | 15 | 1.5 | 2510 | 20 | AAZ41256 | Human normal ovar | 468 | 15 | 1.5 | 5375 | 20 | AAV98274 | Human tumour supp |
| C 396 | 15 | 1.5 | 2517 | 20 | AAZ33560 | Human breast tumou | 469 | 15 | 1.5 | 5531 | 21 | AACT5588 | Human ORFX ORF1143 |
| C 397 | 15 | 1.5 | 2524 | 21 | AACT7206 | Human ORFX ORF2761 | 470 | 15 | 1.5 | 6570 | 20 | AAZ20540 | Poly(nucleotide seq |
| C 398 | 15 | 1.5 | 2526 | 20 | AAO06819 | Chlamydia pneumoni | 471 | 15 | 1.5 | 6850 | 19 | AAV69920 | Expression vector |
| C 399 | 15 | 1.5 | 2526 | 21 | AAAZ8691 | C. pneumoniae CPNI | 472 | 15 | 1.5 | 7989 | 21 | AAV98968 | Hepatitis C virus |
| C 400 | 15 | 1.5 | 2538 | 21 | AAAS5217 | C. symbiosum open | 473 | 15 | 1.5 | 8001 | 21 | AAV98967 | Hepatitis C virus |
| C 401 | 15 | 1.5 | 2551 | 9 | AAAB1251 | Bio F, Bio C and B | 474 | 15 | 1.5 | 8528 | 14 | AAO46249 | Construct EC2L (Co |
| C 402 | 15 | 1.5 | 2623 | 21 | AACT5996 | Human ORFX ORF151 | 475 | 15 | 1.5 | 8637 | 21 | AAV98966 | Hepatitis C virus |
| C 403 | 15 | 1.5 | 2635 | 20 | AAV84521 | Human secreted pro | 476 | 15 | 1.5 | 8649 | 21 | AAV98969 | Hepatitis C virus |
| C 404 | 15 | 1.5 | 2646 | 16 | AAO2405 | Rice mature pullul | 477 | 15 | 1.5 | 8906 | 18 | AAV85392 | Human TRIO phospho |
| C 405 | 15 | 1.5 | 2649 | 20 | AAV37749 | Rat HNK-1 cDNA. R | 478 | 15 | 1.5 | 9558 | 16 | AAO88228 | Valencia orange ri |
| C 406 | 15 | 1.5 | 2653 | 18 | AAAT0174 | Proliferation and | 479 | 15 | 1.5 | 9706 | 21 | AAV65342 | NDO related comple |
| C 407 | 15 | 1.5 | 2663 | 20 | AAAI6300 | Human delta-1 prot | 480 | 15 | 1.5 | 9840 | 21 | AAV65340 | NDO related comple |
| C 408 | 15 | 1.5 | 2663 | 20 | AAAI6817 | Human delta-1 gene | 481 | 15 | 1.5 | 9994 | 22 | AAAC85191 | S. avermitilis 10 |
| C 409 | 15 | 1.5 | 2750 | 21 | AAAB2690 | C. pneumoniae CPNI | 482 | 15 | 1.5 | 10427 | 21 | AAZ36325 | Mechanical stress |
| C 410 | 15 | 1.5 | 2762 | 9 | AAAB0976 | Sequence of 2 cont | 483 | 15 | 1.5 | 10897 | 17 | AAAT09187 | Mutu putative onco |
| C 411 | 15 | 1.5 | 2764 | 18 | AAO84358 | STR7. I gene associ | 484 | 15 | 1.5 | 10898 | 21 | AAAS2462 | Human MN gene. Ho |
| C 412 | 15 | 1.5 | 2768 | 12 | AAO10502 | Poly(hydroxybutyrat | 485 | 15 | 1.5 | 10898 | 21 | AAAI6543 | Human MN complete |
| C 413 | 15 | 1.5 | 2824 | 21 | AAAC59843 | Human secreted pro | 486 | 15 | 1.5 | 11076 | 21 | AAV98965 | Hepatitis C virus |
| C 414 | 15 | 1.5 | 2883 | 18 | AAI58898 | C-Delta-1 gene (al | 487 | 15 | 1.5 | 11279 | 21 | AAV98965 | Hepatitis C virus |
| C 415 | 15 | 1.5 | 2933 | 21 | AAAS4105 | PRO172 cDNA. Homo | 488 | 15 | 1.5 | 11871 | 20 | AAAI13108 | Enterococcus faeca |
| C 416 | 15 | 1.5 | 2933 | 21 | AAAC58587 | Human PRO172 prote | 489 | 15 | 1.5 | 12019 | 20 | AAAI1867 | Alcaligenes sp. Po |
| C 417 | 15 | 1.5 | 2933 | 21 | AAAT7512 | Human PRO172 cDNA | 490 | 15 | 1.5 | 12286 | 21 | AAZ35261 | Plant generic retr |
| C 418 | 15 | 1.5 | 2933 | 21 | AAAC93703 | PRO172 DNA35916-11 | 491 | 15 | 1.5 | 12571 | 21 | AAZ35272 | Soybean retroleme |
| C 419 | 15 | 1.5 | 2933 | 22 | AAAC97368 | Human angiogenesis | 492 | 15 | 1.5 | 12808 | 21 | AAV65347 | NDO related comple |
| C 420 | 15 | 1.5 | 2982 | 16 | AAAT02400 | Pullulanase expres | 493 | 15 | 1.5 | 12886 | 21 | AAO59084 | Human patched-like |
| C 421 | 15 | 1.5 | 2988 | 16 | AAAT02399 | Rice mature pullul | 494 | 15 | 1.5 | 13518 | 20 | AAV20563 | Polynucleotide seq |
| C 422 | 15 | 1.5 | 2999 | 21 | AACT81119 | Human cancer assoc | 495 | 15 | 1.5 | 13574 | 21 | AAAB15293 | N. meningitidis pa |
| C 423 | 15 | 1.5 | 3035 | 21 | AAAF63462 | Streptomyces globi | 496 | 15 | 1.5 | 16020 | 21 | AAV39283 | Streptomyces nogai |
| C 424 | 15 | 1.5 | 3046 | 22 | AAAF5137 | Murine semaphorin | 497 | 15 | 1.5 | 18436 | 19 | AAV52220 | Streptococcus pneu |
| C 425 | 15 | 1.5 | 3046 | 22 | AAAF5140 | Murine DNA: SEQ. ID | 498 | 15 | 1.5 | 24417 | 18 | AAAT97221 | Pseudomonas aerugi |
| C 426 | 15 | 1.5 | 3103 | 21 | AAZ50708 | Nucleotide sequenc | 499 | 15 | 1.5 | 25871 | 21 | AAV98968 | Human genomic OCTN |
| C 427 | 15 | 1.5 | 3105 | 18 | AAV74840 | Staphylococcus aur | 500 | 15 | 1.5 | 26281 | 21 | AAZ60929 | Nucleotide sequenc |
| C 428 | 15 | 1.5 | 3105 | 18 | AAV94468 | Human Fcnd540 gene | 501 | 15 | 1.5 | 26778 | 21 | AAAB1477 | N. meningitidis pa |
| C 429 | 15 | 1.5 | 3111 | 22 | AAAF26660 | Human Smd7 nucleo | 502 | 15 | 1.5 | 32998 | 21 | AAAS5186 | Cenarchaeum symbio |
| C 430 | 15 | 1.5 | 3141 | 14 | AAO49903 | Nicotiana tabacum | 503 | 15 | 1.5 | 36519 | 19 | AAV22141 | Chimpanzee adenovi |
| C 431 | 15 | 1.5 | 3153 | 19 | AAV18995 | Glucocorticoidase ge | 504 | 15 | 1.5 | 37544 | 21 | AAAS50029 | Cosmid cHR15 enco |
| C 432 | 15 | 1.5 | 3162 | 21 | AAZ98679 | Human delta protei | 505 | 15 | 1.5 | 42000 | 21 | AAV65349 | Streptomyces globi |
| C 433 | 15 | 1.5 | 3218 | 22 | AAAF63440 | Human ADAMTS-7 CDN | 506 | 15 | 1.5 | 44576 | 21 | AAZ61522 | Cosmid CV014 conta |
| C 434 | 15 | 1.5 | 3284 | 20 | AAO02677 | T. versicolor lacc | 507 | 15 | 1.5 | 50341 | 19 | AAV22674 | L5 shuttle phasid |
| C 435 | 15 | 1.5 | 3300 | 18 | AAAT62589 | Thermotoga neopoli | 508 | 15 | 1.5 | 50341 | 21 | AAZ39519 | Streptococcus olea |
| C 436 | 15 | 1.5 | 3309 | 21 | AAAF21886 | Human breast and o | 509 | 15 | 1.5 | 50937 | 21 | AAV98969 | Styobacteriophage |
| C 437 | 15 | 1.5 | 3337 | 10 | AAAN91089 | Fragment of Pseudo | 510 | 15 | 1.5 | 52297 | 16 | AAAT51411 | L5 mycobacterioph |
| C 438 | 15 | 1.5 | 3435 | 8 | AAAT0434 | Sequence of S.11v1 | 511 | 15 | 1.5 | 52298 | 14 | AAO47357 | Nucleotide sequenc |
| C 439 | 15 | 1.5 | 3435 | 14 | AAO48429 | Streptomyces livid | 512 | 15 | 1.5 | 58857 | 21 | AAAS5471 | Streptomyces globi |
| C 440 | 15 | 1.5 | 3517 | 17 | AAAT31287 | S. lividans gal op | 513 | 15 | 1.5 | 63164 | 21 | AAAF63348 | WFS1 variant genom |
| C 441 | 15 | 1.5 | 3517 | 17 | AAAT31287 | Rabbit poly-immuno | 514 | 15 | 1.5 | 67212 | 21 | AAAO8954 | N. meningitidis pa |
| C 442 | 15 | 1.5 | 3532 | 21 | AAAS2530 | Human MN 5' UTR. fr | 515 | 15 | 1.5 | 78845 | 21 | AAAB1463 | Genomic fragment # |
| C 443 | 15 | 1.5 | 3532 | 21 | AAAI6611 | Human hippocampal | 516 | 15 | 1.5 | 98629 | 22 | AAV28550 | BAC containing rep |
| C 444 | 15 | 1.5 | 3550 | 20 | AAK99701 | Human hippocampal | 517 | 15 | 1.5 | 109973 | 21 | AAV22298 | Nesleria meningit |
| C 445 | 15 | 1.5 | 3550 | 22 | AAAF24125 | Human hippocampal | 518 | 15 | 1.5 | 172325 | 21 | AAAF21613 | Hereditary haemoch |
| C 446 | 15 | 1.5 | 3571 | 20 | AAV99702 | Human mammary sel- | 519 | 15 | 1.5 | 235033 | 19 | AAV57926 | Hereditary haemoch |
| C 447 | 15 | 1.5 | 3571 | 22 | AAAF24126 | Human mammary sel- | 520 | 15 | 1.5 | 237326 | 19 | AAV57903 | Neisseria meningit |
| C 448 | 15 | 1.5 | 3821 | 15 | AAO56609 | Human PGE-2/MF-3 5 | 521 | 15 | 1.5 | 349980 | 21 | AAV21608 | Neisseria meningit |
| C 449 | 15 | 1.5 | 3821 | 20 | AAK34365 | Human nerve growth | 522 | 15 | 1.5 | 349980 | 21 | AAV21612 | Neisseria meningit |

| | | | | | | | | | | | | | |
|-----|----|-----|---------|-----|-----------|--------------------|-----|----|-----|-----|----|-----------|---------------------|
| 523 | 15 | 1.5 | 1082138 | 21 | AAE22305 | Arabidopsis thalia | 596 | 14 | 1.4 | 246 | 21 | AAC09887 | Human secreted pro |
| 524 | 15 | 1.5 | 1230025 | 20 | AAAX91990 | Nucleotide sequenc | 597 | 14 | 1.4 | 251 | 19 | AAAX12270 | Human biallelic po |
| 525 | 14 | 1.4 | | 15 | AAZ59280 | Human NR8 gene pro | 598 | 14 | 1.4 | 255 | 11 | AAO03507 | Sequence encoding |
| 526 | 14 | 1.4 | | 15 | AAE47941 | IGFBP3 oligonucleo | 599 | 14 | 1.4 | 255 | 18 | AAE88164 | Nucleotide sequenc |
| 527 | 14 | 1.4 | | 15 | AAE47942 | IGFBP3 oligonucleo | 600 | 14 | 1.4 | 255 | 18 | AAE94572 | Polynucleotide seq |
| 528 | 14 | 1.4 | | 18 | AAZ32752 | Human protease-act | 601 | 14 | 1.4 | 255 | 21 | AAAS0098 | Arabidopsis herbic |
| 529 | 14 | 1.4 | | 20 | AAO62027 | Mutant Ki-ras 5'-U | 602 | 14 | 1.4 | 255 | 21 | AAC18419 | Human secreted pro |
| 530 | 14 | 1.4 | | 20 | AAO79846 | K-ras modulating s | 603 | 14 | 1.4 | 263 | 21 | AAC15554 | Human secreted pro |
| 531 | 14 | 1.4 | | 20 | AAAX96704 | PCR primer used to | 604 | 14 | 1.4 | 265 | 21 | AAE75473 | Partial cDNA seque |
| 532 | 14 | 1.4 | | 20 | AAAS6986 | Ras gene modulator | 605 | 14 | 1.4 | 265 | 21 | AAC31308 | Human secreted pro |
| 533 | 14 | 1.4 | | 20 | AAAX21622 | Human Ki-ras speci | 606 | 14 | 1.4 | 268 | 21 | AAC02717 | Human secreted pro |
| 534 | 14 | 1.4 | | 20 | AAAS4026 | Antisense oligonuc | 607 | 14 | 1.4 | 269 | 21 | AAC16768 | Human secreted pro |
| 535 | 14 | 1.4 | | 20 | AAAS5860 | Human Ki-ras antis | 608 | 14 | 1.4 | 270 | 18 | AAE67889 | Human secreted pro |
| 536 | 14 | 1.4 | | 21 | AAAX09193 | Human biallelic po | 609 | 14 | 1.4 | 270 | 20 | AAC74427 | Human secreted pro |
| 537 | 14 | 1.4 | | 27 | AAE64537 | PCR primer G7 used | 610 | 14 | 1.4 | 276 | 19 | AAV20119 | Human secreted pro |
| 538 | 14 | 1.4 | | 29 | AAAI9247 | Integrin alpha 6 s | 611 | 14 | 1.4 | 282 | 11 | AAO03335 | Probe (20) for mic |
| 539 | 14 | 1.4 | | 29 | AAAI9772 | Integrin alpha 6 s | 612 | 14 | 1.4 | 282 | 14 | AAQ42653 | Fragment of post-l |
| 540 | 14 | 1.4 | | 30 | AAAI8336 | Human cytochrome P | 613 | 14 | 1.4 | 291 | 21 | AAC00592 | pt-NAB virus lamb |
| 541 | 14 | 1.4 | | 30 | AAE48336 | Primer for 0.6 kb | 614 | 14 | 1.4 | 293 | 20 | AAAS1660 | Human secreted pro |
| 542 | 14 | 1.4 | | 30 | AAE26928 | Human cytochrome P | 615 | 14 | 1.4 | 293 | 20 | AAV87616 | EST clone BC259. |
| 543 | 14 | 1.4 | | 31 | AAE17403 | Human derived cyto | 616 | 14 | 1.4 | 294 | 21 | AAC05481 | Human secreted pro |
| 544 | 14 | 1.4 | | 31 | AAE67562 | Nucleotide fragmen | 617 | 14 | 1.4 | 295 | 15 | AAQ73711 | Retinotansposon pro |
| 545 | 14 | 1.4 | | 32 | AAE59008 | Mouse epididymis-s | 618 | 14 | 1.4 | 295 | 21 | AAC06337 | Human secreted pro |
| 546 | 14 | 1.4 | | 37 | AAE03355 | Anti-IL-8 Mab 5.12 | 619 | 14 | 1.4 | 297 | 17 | AAE09205 | Virulence factor s |
| 547 | 14 | 1.4 | | 37 | AAV03191 | Anti IL-8 antibody | 620 | 14 | 1.4 | 298 | 17 | AAE09207 | Human gene express |
| 548 | 14 | 1.4 | | 37 | AAE93560 | Murine Mab (5.12.1 | 621 | 14 | 1.4 | 300 | 20 | AAZ12673 | Human gene express |
| 549 | 14 | 1.4 | | 37 | AAE78555 | Monoclonal antibod | 622 | 14 | 1.4 | 300 | 20 | AAZ12605 | Human gene express |
| 550 | 14 | 1.4 | | 37 | AAE55083 | Murine anti-IL-8 M | 623 | 14 | 1.4 | 300 | 21 | AAAO0198 | Human colon cancer |
| 551 | 14 | 1.4 | | 37 | AAV10277 | Murine monoclonal | 624 | 14 | 1.4 | 300 | 21 | AAAO0442 | Human colon cancer |
| 552 | 14 | 1.4 | | 37 | AAV06432 | Murine Mab (5.12.1 | 625 | 14 | 1.4 | 300 | 21 | AAAO1144 | Human colon cancer |
| 553 | 14 | 1.4 | | 37 | AAE90552 | Murine 5.12.14 ant | 626 | 14 | 1.4 | 303 | 21 | AAAO1374 | Human colon cancer |
| 554 | 14 | 1.4 | | 37 | AAE65479 | Murine anti-IL-8 a | 627 | 14 | 1.4 | 304 | 21 | AAE18914 | Human colon cancer |
| 555 | 14 | 1.4 | | 37 | AAE287940 | Murine anti-IL-8 a | 628 | 14 | 1.4 | 304 | 21 | AAC30331 | DNA encoding spide |
| 556 | 14 | 1.4 | | 40 | AAE286124 | Polynucleotide seq | 629 | 14 | 1.4 | 304 | 21 | AAE242955 | Human secreted pro |
| 557 | 14 | 1.4 | | 42 | AAE05507 | Probe to N-termina | 630 | 14 | 1.4 | 304 | 21 | AAE242955 | Human 5' EST Isola |
| 558 | 14 | 1.4 | | 42 | AAO15219 | Human fibronectin | 631 | 14 | 1.4 | 305 | 18 | AAV12653 | VH251/gammal mutan |
| 559 | 14 | 1.4 | | 47 | AAE28991 | Sequence of PCR pr | 632 | 14 | 1.4 | 305 | 19 | AAE66013 | Human gene signatu |
| 560 | 14 | 1.4 | | 48 | AAV19184 | Human growth facto | 633 | 14 | 1.4 | 312 | 16 | AAE24501 | Nucleotide sequenc |
| 561 | 14 | 1.4 | | 48 | AAE63329 | PCR primer specifi | 634 | 14 | 1.4 | 313 | 16 | AAE19149 | Human gene signatu |
| 562 | 14 | 1.4 | | 54 | AAE26611 | Polyketide synthas | 635 | 14 | 1.4 | 314 | 21 | AAC23276 | Human secreted pro |
| 563 | 14 | 1.4 | | 90 | AAV17330 | Primer CME00972 us | 636 | 14 | 1.4 | 316 | 21 | AAC00622 | Human secreted pro |
| 564 | 14 | 1.4 | | 97 | AAE12931 | Human secreted pro | 637 | 14 | 1.4 | 316 | 21 | AAZ42333 | Human 5' EST Isola |
| 565 | 14 | 1.4 | | 101 | AAV02808 | Human RHAM genom | 638 | 14 | 1.4 | 320 | 21 | AAC01859 | Human secreted pro |
| 566 | 14 | 1.4 | | 108 | AAE90209 | Recombinant DNA fo | 639 | 14 | 1.4 | 322 | 21 | AAC74961 | Human ORFX ORF51 |
| 567 | 14 | 1.4 | | 110 | AAE29865 | Human secreted pro | 640 | 14 | 1.4 | 324 | 21 | AAE71620 | N. meningitidis BA |
| 568 | 14 | 1.4 | | 115 | AAE21334 | Human gene signatu | 641 | 14 | 1.4 | 324 | 21 | AAE53434 | Neisseria meningit |
| 569 | 14 | 1.4 | | 117 | AAQ26958 | Human gene signatu | 642 | 14 | 1.4 | 324 | 21 | AAE53435 | Neisseria meningit |
| 570 | 14 | 1.4 | | 124 | AAE21511 | Human biallelic po | 643 | 14 | 1.4 | 326 | 22 | AAE63340 | Human XAG growth f |
| 571 | 14 | 1.4 | | 132 | AAE10479 | Human biallelic po | 644 | 14 | 1.4 | 326 | 22 | AAE63340 | HuXAG-2 related cd |
| 572 | 14 | 1.4 | | 138 | AAE32093 | Human secreted pro | 645 | 14 | 1.4 | 328 | 21 | AAC22609 | Human secreted pro |
| 573 | 14 | 1.4 | | 153 | AAE73766 | Brevibacterium fla | 646 | 14 | 1.4 | 330 | 20 | AAE89310 | EST clone CM423. |
| 574 | 14 | 1.4 | | 153 | AAE6761 | Seq ID No: 1 of JP | 647 | 14 | 1.4 | 331 | 20 | AAE219430 | M. tuberculosis an |
| 575 | 14 | 1.4 | | 166 | AAE37721 | Human PRO266 clone | 648 | 14 | 1.4 | 331 | 20 | AAE219218 | M. tuberculosis re |
| 576 | 14 | 1.4 | | 171 | AAE42248 | Human secreted exp | 649 | 14 | 1.4 | 332 | 21 | AAE08918 | Fusarium venenatum |
| 577 | 14 | 1.4 | | 172 | AAE31926 | Human secreted pro | 650 | 14 | 1.4 | 335 | 21 | AAC70608 | Single nucleotide |
| 578 | 14 | 1.4 | | 177 | AAE90208 | Staphylococcus pro | 651 | 14 | 1.4 | 338 | 21 | AAC14174 | Human secreted pro |
| 579 | 14 | 1.4 | | 177 | AAO11649 | FB fragment of pro | 652 | 14 | 1.4 | 341 | 21 | AAE11086 | Fusarium venenatum |
| 580 | 14 | 1.4 | | 192 | AAO01767 | Binding region for | 653 | 14 | 1.4 | 343 | 20 | AAE28036 | Partial sequence o |
| 581 | 14 | 1.4 | | 202 | AAE2960 | Human secreted exp | 654 | 14 | 1.4 | 346 | 21 | AAE15981 | Human secreted pro |
| 582 | 14 | 1.4 | | 209 | AAE57088 | Human colon cancer | 655 | 14 | 1.4 | 348 | 13 | AAE28981 | Sequence encoding |
| 583 | 14 | 1.4 | | 213 | AAE51795 | Human secreted pro | 656 | 14 | 1.4 | 353 | 20 | AAE66600 | EST clone BE101. |
| 584 | 14 | 1.4 | | 215 | AAO10349 | Encodes Immunoglob | 657 | 14 | 1.4 | 353 | 21 | AAC75340 | Human ORFX ORF95 |
| 585 | 14 | 1.4 | | 216 | AAE51573 | Human secreted pro | 658 | 14 | 1.4 | 354 | 22 | AAE66171 | Novel human polyu |
| 586 | 14 | 1.4 | | 218 | AAO76820 | Human genome fragm | 659 | 14 | 1.4 | 357 | 13 | AAE28982 | Sequence encoding |
| 587 | 14 | 1.4 | | 218 | AAE67074 | Eucalyptus grandis | 660 | 14 | 1.4 | 357 | 21 | AAC06032 | Human secreted pro |
| 588 | 14 | 1.4 | | 220 | AAE41733 | Human secreted exp | 661 | 14 | 1.4 | 357 | 21 | AAC09562 | Human secreted pro |
| 589 | 14 | 1.4 | | 227 | AAE90320 | EST clone DK230. | 662 | 14 | 1.4 | 360 | 18 | AAV17306 | Humanised antibody |
| 590 | 14 | 1.4 | | 227 | AAE15874 | DNA encoding VEGF- | 663 | 14 | 1.4 | 360 | 18 | AAV17307 | Humanised antibody |
| 591 | 14 | 1.4 | | 228 | AAE08265 | Human secreted pro | 664 | 14 | 1.4 | 360 | 18 | AAV17308 | Humanised antibody |
| 592 | 14 | 1.4 | | 231 | AAC02414 | Human secreted pro | 665 | 14 | 1.4 | 360 | 18 | AAV17309 | Humanised antibody |
| 593 | 14 | 1.4 | | 240 | AAE5346 | Myrtaceae microsat | 666 | 14 | 1.4 | 360 | 18 | AAV17310 | Humanised antibody |
| 594 | 14 | 1.4 | | 242 | AAE88009 | Human T gene DNA f | 667 | 14 | 1.4 | 360 | 18 | AAV17311 | Humanised antibody |
| 595 | 14 | 1.4 | | 242 | AAC08117 | Human secreted pro | 668 | 14 | 1.4 | 360 | 18 | AAV17271 | Anti-CEA antibody |

| | | | | | | | | | | | | | |
|-------|----|-----|-----|----|----------|--------------------|-------|----|-----|-----|----|----------|---------------------|
| 669 | 14 | 1.4 | 363 | 21 | AA254189 | Neisseria meningit | c 742 | 14 | 1.4 | 446 | 21 | AA20933 | Human secreted pro |
| 670 | 14 | 1.4 | 363 | 21 | AA254190 | Neisseria meningit | 743 | 14 | 1.4 | 448 | 18 | AA258802 | Human telomerase R |
| c 671 | 14 | 1.4 | 363 | 22 | AA254195 | Human 10 kd Clara | 744 | 14 | 1.4 | 448 | 21 | AA274869 | Human ORFX ORF424 |
| c 672 | 14 | 1.4 | 363 | 22 | AA254196 | Human 10 kd Clara | c 745 | 14 | 1.4 | 448 | 21 | AA257415 | Archidonic acid m |
| c 673 | 14 | 1.4 | 363 | 22 | AA254197 | Human 10 kd Clara | c 746 | 14 | 1.4 | 448 | 21 | AA257416 | Archidonic acid m |
| c 674 | 14 | 1.4 | 363 | 22 | AA254198 | Human 10 kd Clara | c 747 | 14 | 1.4 | 448 | 21 | AA257417 | Archidonic acid m |
| c 675 | 14 | 1.4 | 363 | 22 | AA254199 | Human 10 kd Clara | c 748 | 14 | 1.4 | 448 | 21 | AA257418 | Archidonic acid m |
| c 676 | 14 | 1.4 | 363 | 22 | AA254200 | Human 10 kd Clara | c 749 | 14 | 1.4 | 449 | 20 | AA257419 | Archidonic acid m |
| c 677 | 14 | 1.4 | 366 | 21 | AA254201 | Human secreted pro | c 750 | 14 | 1.4 | 451 | 19 | AA257420 | EST clone CT729 |
| c 678 | 14 | 1.4 | 367 | 22 | AA254202 | Human Clara cell p | 751 | 14 | 1.4 | 452 | 21 | AA257421 | RNA component of h |
| c 679 | 14 | 1.4 | 369 | 13 | AA254203 | Sequence encoding | 752 | 14 | 1.4 | 454 | 14 | AA257422 | Human secreted pro |
| c 680 | 14 | 1.4 | 374 | 21 | AA254204 | Human secreted pro | 753 | 14 | 1.4 | 456 | 22 | AA257423 | Human brain Expres |
| c 681 | 14 | 1.4 | 375 | 21 | AA254205 | Human secreted pro | c 754 | 14 | 1.4 | 458 | 22 | AA257424 | C. glutamicum SRT |
| c 682 | 14 | 1.4 | 377 | 22 | AA254206 | Archidopsis thalia | c 755 | 14 | 1.4 | 461 | 21 | AA257425 | Human secreted pro |
| c 683 | 14 | 1.4 | 377 | 22 | AA254207 | Human 10 kd Clara | c 756 | 14 | 1.4 | 464 | 21 | AA257426 | Human secreted pro |
| c 684 | 14 | 1.4 | 377 | 22 | AA254208 | Human 10 kd Clara | c 757 | 14 | 1.4 | 466 | 21 | AA257427 | Plus radiata sur |
| c 685 | 14 | 1.4 | 378 | 22 | AA254209 | Human 10 kd Clara | c 758 | 14 | 1.4 | 466 | 22 | AA257428 | Zea mays DNA fragm |
| c 686 | 14 | 1.4 | 378 | 22 | AA254210 | M tuberculosis Rv3 | c 759 | 14 | 1.4 | 468 | 22 | AA257429 | Novel human polynu |
| c 687 | 14 | 1.4 | 378 | 22 | AA254211 | Human 10 kd Clara | c 760 | 14 | 1.4 | 472 | 21 | AA257430 | Archidopsis thalia |
| c 688 | 14 | 1.4 | 378 | 22 | AA254212 | Human 10 kd Clara | c 761 | 14 | 1.4 | 473 | 21 | AA257431 | Human secreted pro |
| c 689 | 14 | 1.4 | 381 | 14 | AA254213 | Human brain Expres | c 762 | 14 | 1.4 | 475 | 21 | AA257432 | Human secreted pro |
| c 690 | 14 | 1.4 | 381 | 19 | AA254214 | Neisseria meningit | c 763 | 14 | 1.4 | 475 | 21 | AA257433 | Zea mays DNA fragm |
| c 691 | 14 | 1.4 | 381 | 21 | AA254215 | Human secreted pro | c 764 | 14 | 1.4 | 477 | 21 | AA257434 | Human secreted pro |
| c 692 | 14 | 1.4 | 384 | 22 | AA254216 | Novel human polynu | c 765 | 14 | 1.4 | 478 | 21 | AA257435 | Human ORFX ORF1742 |
| c 693 | 14 | 1.4 | 385 | 18 | AA254217 | Humanised anti-CEA | c 766 | 14 | 1.4 | 481 | 13 | AA257436 | Human secreted pro |
| c 694 | 14 | 1.4 | 386 | 9 | AA254218 | Humanised anti-CEA | c 767 | 14 | 1.4 | 481 | 18 | AA257437 | Phenolone receptor |
| c 695 | 14 | 1.4 | 386 | 9 | AA254219 | DNA encoding modif | c 768 | 14 | 1.4 | 481 | 18 | AA257438 | Staphylococcus aur |
| c 696 | 14 | 1.4 | 387 | 9 | AA254220 | DNA encoding biosy | c 769 | 14 | 1.4 | 485 | 20 | AA257439 | Nucleic acid E69r |
| c 697 | 14 | 1.4 | 391 | 22 | AA254221 | Novel human polynu | c 770 | 14 | 1.4 | 487 | 21 | AA257440 | Human secreted pro |
| c 698 | 14 | 1.4 | 394 | 22 | AA254222 | Novel human polynu | c 771 | 14 | 1.4 | 488 | 21 | AA257441 | Zea mays DNA fragm |
| c 699 | 14 | 1.4 | 395 | 21 | AA254223 | Human secreted pro | c 772 | 14 | 1.4 | 489 | 20 | AA257442 | Potpholyonas gling |
| c 700 | 14 | 1.4 | 395 | 21 | AA254224 | Human 10 kd Clara | c 773 | 14 | 1.4 | 490 | 21 | AA257443 | Human telomerase R |
| c 701 | 14 | 1.4 | 396 | 21 | AA254225 | DNA encoding light | c 774 | 14 | 1.4 | 491 | 14 | AA257444 | Expressed Sequence |
| c 702 | 14 | 1.4 | 396 | 21 | AA254226 | Human secreted pro | c 775 | 14 | 1.4 | 491 | 14 | AA257445 | Human brain Expres |
| c 703 | 14 | 1.4 | 396 | 22 | AA254227 | Human ovarian can | c 776 | 14 | 1.4 | 492 | 21 | AA257446 | Human ORFX ORF1621 |
| c 704 | 14 | 1.4 | 397 | 20 | AA254228 | EST clone B197. H | c 777 | 14 | 1.4 | 492 | 21 | AA257447 | N. meningitidis pa |
| c 705 | 14 | 1.4 | 398 | 14 | AA254229 | Human brain Expres | c 778 | 14 | 1.4 | 492 | 22 | AA257448 | Nucleotide sequenc |
| c 706 | 14 | 1.4 | 399 | 17 | AA254230 | Partial sequence o | c 779 | 14 | 1.4 | 500 | 21 | AA257449 | Zea mays DNA fragm |
| c 707 | 14 | 1.4 | 399 | 17 | AA254231 | Partial sequence o | c 780 | 14 | 1.4 | 501 | 17 | AA257450 | Immunogen DNA fragm |
| c 708 | 14 | 1.4 | 399 | 17 | AA254232 | Partial sequence o | c 781 | 14 | 1.4 | 505 | 19 | AA257451 | Human XAS growth f |
| c 709 | 14 | 1.4 | 399 | 17 | AA254233 | Partial sequence o | c 782 | 14 | 1.4 | 505 | 20 | AA257452 | Human CAS847 deriv |
| c 710 | 14 | 1.4 | 400 | 10 | AA254234 | Fc binding region | c 783 | 14 | 1.4 | 505 | 22 | AA257453 | HuXAG-2 related cd |
| c 711 | 14 | 1.4 | 400 | 18 | AA254235 | Staphylococcus aur | c 784 | 14 | 1.4 | 507 | 20 | AA257454 | Corn CDC-16 cDNA c |
| c 712 | 14 | 1.4 | 402 | 21 | AA254236 | Human ORFX ORF558 | c 785 | 14 | 1.4 | 507 | 21 | AA257455 | Archidopsis thalia |
| c 713 | 14 | 1.4 | 403 | 18 | AA254237 | Staphylococcus aur | c 786 | 14 | 1.4 | 508 | 21 | AA257456 | Archidopsis thalia |
| c 714 | 14 | 1.4 | 403 | 22 | AA254238 | Novel human polynu | c 787 | 14 | 1.4 | 510 | 20 | AA257457 | Porphyromonas gling |
| c 715 | 14 | 1.4 | 404 | 21 | AA254239 | Nucleotide sequenc | c 788 | 14 | 1.4 | 510 | 21 | AA257458 | DNA encoding a pol |
| c 716 | 14 | 1.4 | 405 | 20 | AA254240 | Human anti-angioge | c 789 | 14 | 1.4 | 512 | 21 | AA257459 | Human MN gene intr |
| c 717 | 14 | 1.4 | 406 | 21 | AA254241 | Human secreted pro | c 790 | 14 | 1.4 | 512 | 21 | AA257460 | Human MN intron 7 |
| c 718 | 14 | 1.4 | 407 | 14 | AA254242 | Human brain Expres | c 791 | 14 | 1.4 | 513 | 17 | AA257461 | HIV-1 group O stra |
| c 719 | 14 | 1.4 | 408 | 20 | AA254243 | Nucleotide sequenc | c 792 | 14 | 1.4 | 513 | 21 | AA257462 | Fusarium venenatum |
| c 720 | 14 | 1.4 | 408 | 20 | AA254244 | Nucleotide sequenc | c 793 | 14 | 1.4 | 513 | 21 | AA257463 | Archidopsis thalia |
| c 721 | 14 | 1.4 | 409 | 21 | AA254245 | Human ORFX ORF206 | c 794 | 14 | 1.4 | 516 | 20 | AA257464 | Clone HP01766 enco |
| c 722 | 14 | 1.4 | 409 | 22 | AA254246 | Novel human polynu | c 795 | 14 | 1.4 | 516 | 22 | AA257465 | Murine TANGO 273 O |
| c 723 | 14 | 1.4 | 412 | 19 | AA254247 | Targeting vector T | c 796 | 14 | 1.4 | 518 | 21 | AA257466 | Trichoderma reesei |
| c 724 | 14 | 1.4 | 412 | 21 | AA254248 | Fusarium venenatum | c 797 | 14 | 1.4 | 520 | 22 | AA257467 | Human secreted pro |
| c 725 | 14 | 1.4 | 418 | 18 | AA254249 | Mouse GDNFR alpha | c 798 | 14 | 1.4 | 522 | 20 | AA257468 | S. aureus phosphor |
| c 726 | 14 | 1.4 | 418 | 21 | AA254250 | Human secreted exp | c 799 | 14 | 1.4 | 522 | 21 | AA257469 | L. esculentum MAP |
| c 727 | 14 | 1.4 | 420 | 21 | AA254251 | Human secreted pro | c 800 | 14 | 1.4 | 522 | 22 | AA257470 | Korean mistletoe 1 |
| c 728 | 14 | 1.4 | 421 | 14 | AA254252 | Human glucose regu | c 801 | 14 | 1.4 | 523 | 21 | AA257471 | Mouse F1222 (inhib |
| c 729 | 14 | 1.4 | 421 | 20 | AA254253 | EST clone DD285 | c 802 | 14 | 1.4 | 524 | 21 | AA257472 | Trichoderma reesei |
| c 730 | 14 | 1.4 | 421 | 21 | AA254254 | Human ORFX ORF2179 | c 803 | 14 | 1.4 | 525 | 20 | AA257473 | S. aureus phosphor |
| c 731 | 14 | 1.4 | 427 | 22 | AA254255 | Human lung tumour | c 804 | 14 | 1.4 | 525 | 21 | AA257474 | Fusarium venenatum |
| c 732 | 14 | 1.4 | 433 | 21 | AA254256 | Lung cancer associ | c 805 | 14 | 1.4 | 526 | 22 | AA257475 | CDNA encoding CD1E |
| c 733 | 14 | 1.4 | 433 | 21 | AA254257 | E. coli prolifera | c 806 | 14 | 1.4 | 527 | 22 | AA257476 | Testis Enhanced Ge |
| c 734 | 14 | 1.4 | 435 | 11 | AA254258 | Clone hps3 encodin | c 807 | 14 | 1.4 | 528 | 21 | AA257477 | Human epidermal pr |
| c 735 | 14 | 1.4 | 436 | 20 | AA254259 | EST clone CT738 | c 808 | 14 | 1.4 | 528 | 21 | AA257478 | DNA encoding a mar |
| c 736 | 14 | 1.4 | 438 | 21 | AA254260 | Archidopsis thalia | c 809 | 14 | 1.4 | 534 | 14 | AA254261 | Human glucokinase |
| c 737 | 14 | 1.4 | 438 | 21 | AA254261 | E. coli prolifera | c 810 | 14 | 1.4 | 534 | 14 | AA254262 | Archidopsis thalia |
| c 738 | 14 | 1.4 | 441 | 21 | AA254262 | Eucalyptus grandis | c 811 | 14 | 1.4 | 535 | 19 | AA254263 | DNA encoding a krl |
| c 739 | 14 | 1.4 | 441 | 21 | AA254263 | Human secreted pro | c 812 | 14 | 1.4 | 535 | 21 | AA254264 | White shrimp multi |
| c 740 | 14 | 1.4 | 443 | 21 | AA254264 | Eucalyptus grandis | c 813 | 14 | 1.4 | 536 | 19 | AA254265 | DNA encoding a krl |
| c 741 | 14 | 1.4 | 444 | 21 | AA254265 | Human telomerase R | c 814 | 14 | 1.4 | 536 | 21 | AA254266 | Eucalyptus grandis |

| | | | | | | | | | | | | |
|-------|----|-----|----|-----------|---------------------|-------|----|-----|-----|----|-----------|---------------------|
| C 815 | 14 | 536 | 21 | AAA90471 | White shrimp multi | 888 | 14 | 1.4 | 665 | 21 | AAC03822 | Human secreted pro |
| C 816 | 14 | 538 | 19 | AAV13839 | Homo sapiens telom | C 889 | 14 | 1.4 | 670 | 22 | AAF23410 | Human SRC1 DNA seq |
| C 817 | 14 | 538 | 19 | AAV16093 | Human telomerase p | C 890 | 14 | 1.4 | 672 | 21 | AAC34816 | Arabidopsis thalia |
| C 818 | 14 | 540 | 19 | AAV13833 | Homo sapiens telom | C 891 | 14 | 1.4 | 673 | 21 | AAAO2035 | Human colon cancer |
| C 819 | 14 | 540 | 19 | AAV13838 | Homo sapiens telom | C 892 | 14 | 1.4 | 674 | 21 | AAV12652 | Aspergillus oryzae |
| C 820 | 14 | 540 | 19 | AAV13837 | Homo sapiens telom | C 893 | 14 | 1.4 | 676 | 21 | AAV12612 | Aspergillus oryzae |
| C 821 | 14 | 540 | 19 | AAV05373 | Human telomerase p | C 894 | 14 | 1.4 | 680 | 18 | AAV58803 | Human telomerase e |
| C 822 | 14 | 540 | 19 | AAV05374 | Human telomerase p | C 895 | 14 | 1.4 | 683 | 21 | AAV58147 | Aspergillus oryzae |
| C 823 | 14 | 540 | 19 | AAV16092 | Human telomerase p | C 896 | 14 | 1.4 | 689 | 20 | AAV13926 | Aspergillus oryzae |
| C 824 | 14 | 545 | 21 | AAV09347 | Fusarium venenatum | C 897 | 14 | 1.4 | 689 | 20 | AAV56152 | Human alpha-7 nico |
| C 825 | 14 | 549 | 16 | AAO84578 | Pseudomonas aerugi | C 898 | 14 | 1.4 | 689 | 20 | AAV22655 | Human secreted pro |
| C 826 | 14 | 549 | 21 | AAA45041 | Mouse secreted exp | C 899 | 14 | 1.4 | 695 | 21 | AAC81731 | Human secreted pro |
| C 827 | 14 | 550 | 21 | AAC3988 | Cat flea hindgut a | C 900 | 14 | 1.4 | 700 | 20 | AAZ42018 | Human endometrium |
| C 828 | 14 | 552 | 21 | AAC3988 | Trichoderma reesei | C 901 | 14 | 1.4 | 703 | 20 | AAV84514 | Human secreted pro |
| C 829 | 14 | 556 | 21 | AAA4939 | Human secreted exp | C 902 | 14 | 1.4 | 705 | 21 | AAC51968 | Arabidopsis thalia |
| C 830 | 14 | 559 | 17 | AAV10282 | RNA component of h | C 903 | 14 | 1.4 | 705 | 21 | AAV10094 | Yeast codon-biased |
| C 831 | 14 | 559 | 17 | AAV10282 | Human telomerase R | C 904 | 14 | 1.4 | 705 | 22 | AAV59010 | Mouse epididymis-s |
| C 832 | 14 | 566 | 21 | AAV1031 | Human cancer assoc | C 905 | 14 | 1.4 | 710 | 21 | AAC44000 | Arabidopsis thalia |
| C 833 | 14 | 567 | 20 | AAV09340 | Human Dlnp composi | C 906 | 14 | 1.4 | 714 | 16 | AAO76262 | PNA lectin. Arach |
| C 834 | 14 | 568 | 21 | AAV16436 | Human prostate can | C 907 | 14 | 1.4 | 718 | 21 | AAV14570 | Aspergillus oryzae |
| C 835 | 14 | 572 | 21 | AAV08337 | Fusarium venenatum | C 908 | 14 | 1.4 | 719 | 20 | AAV39664 | Renal cancer assoc |
| C 836 | 14 | 573 | 21 | AAV08762 | Fusarium venenatum | C 909 | 14 | 1.4 | 722 | 22 | AAV22677 | Human gastric can |
| C 837 | 14 | 574 | 20 | AAV90119 | EST Clone DG82. H | C 910 | 14 | 1.4 | 723 | 21 | AAV297462 | Human prostate can |
| C 838 | 14 | 575 | 21 | AAC78221 | Human cancer assoc | C 911 | 14 | 1.4 | 727 | 19 | AAV42211 | Nucleotide sequenc |
| C 839 | 14 | 579 | 20 | AAV01708 | Human anti-angioge | C 912 | 14 | 1.4 | 735 | 21 | AAV02383 | Human colon cancer |
| C 840 | 14 | 591 | 20 | AAV21217 | Neisseria gonorrhoe | C 913 | 14 | 1.4 | 735 | 21 | AAV02383 | Human colon cancer |
| C 841 | 14 | 591 | 21 | AAV254405 | Fusarium venenatum | C 914 | 14 | 1.4 | 738 | 22 | AAV26406 | Pseudomonas sp hea |
| C 842 | 14 | 593 | 21 | AAV09085 | Human prostate can | C 915 | 14 | 1.4 | 743 | 21 | AAC56031 | Eucalyptus grandis |
| C 843 | 14 | 597 | 21 | AAV16408 | Human cancer assoc | C 916 | 14 | 1.4 | 743 | 21 | AAC32823 | Arabidopsis thalia |
| C 844 | 14 | 597 | 21 | AAC77643 | Human prostate can | C 917 | 14 | 1.4 | 744 | 21 | AAV297461 | Human prostate can |
| C 845 | 14 | 597 | 21 | AAA30451 | Human cancer assoc | C 918 | 14 | 1.4 | 746 | 21 | AAC43849 | Arabidopsis thalia |
| C 846 | 14 | 597 | 21 | AAA15654 | HTLV p24 g1 fragme | C 919 | 14 | 1.4 | 750 | 21 | AAV25507 | Canine mature Flt- |
| C 847 | 14 | 599 | 17 | AAV08608 | Protoplasmic membr | C 920 | 14 | 1.4 | 750 | 21 | AAV25508 | Canine mature Flt- |
| C 848 | 14 | 599 | 19 | AAV22055 | DNA encoding a Krt | C 921 | 14 | 1.4 | 755 | 21 | AAC43985 | Arabidopsis thalia |
| C 849 | 14 | 599 | 21 | AAA50470 | White shrimp multi | C 922 | 14 | 1.4 | 756 | 19 | AAV35555 | Human tumour anti |
| C 850 | 14 | 600 | 21 | AAC46771 | Arabidopsis thalia | C 923 | 14 | 1.4 | 757 | 21 | AAC43961 | Arabidopsis thalia |
| C 851 | 14 | 605 | 18 | AAV30888 | Streptococcus pneu | C 924 | 14 | 1.4 | 759 | 20 | AAV38926 | Human validated ca |
| C 852 | 14 | 608 | 21 | AAC49033 | Arabidopsis thalia | C 925 | 14 | 1.4 | 762 | 22 | AAC85472 | A-chain gene isofo |
| C 853 | 14 | 611 | 21 | AAV07754 | Fusarium venenatum | C 926 | 14 | 1.4 | 762 | 22 | AAC85473 | Humanised antibody |
| C 854 | 14 | 611 | 21 | AAC35687 | Arabidopsis thalia | C 927 | 14 | 1.4 | 765 | 18 | AAV17299 | Humanised antibody |
| C 855 | 14 | 616 | 21 | AAV09569 | Fusarium venenatum | C 928 | 14 | 1.4 | 766 | 19 | AAV23103 | CDNA encoding a hu |
| C 856 | 14 | 619 | 21 | AAV280495 | Human colon cancer | C 929 | 14 | 1.4 | 776 | 20 | AAV89573 | Human apoptosis-re |
| C 857 | 14 | 620 | 18 | AAV79540 | Upstream DNA seque | C 930 | 14 | 1.4 | 777 | 21 | AAV54468 | CDNA encoding a ra |
| C 858 | 14 | 623 | 21 | AAV87734 | Human secreted pro | C 931 | 14 | 1.4 | 777 | 21 | AAV08846 | Human DRK-2 DNA op |
| C 859 | 14 | 623 | 21 | AAC33194 | Arabidopsis thalia | C 932 | 14 | 1.4 | 779 | 21 | AAC99883 | Human secreted pro |
| C 860 | 14 | 623 | 22 | AAV64016 | CDNA encoding huma | C 933 | 14 | 1.4 | 780 | 18 | AAV17313 | Humanised heavy ch |
| C 861 | 14 | 624 | 8 | AAV71272 | Sequence of adult | C 934 | 14 | 1.4 | 780 | 22 | AAV22462 | Human breast cance |
| C 862 | 14 | 624 | 21 | AAC57319 | Pinus radiata tran | C 935 | 14 | 1.4 | 785 | 21 | AAV15932 | Human prostate can |
| C 863 | 14 | 626 | 21 | AAC41611 | Arabidopsis thalia | C 936 | 14 | 1.4 | 785 | 22 | AAV72778 | Human prostate can |
| C 864 | 14 | 632 | 21 | AAV14568 | Aspergillus oryzae | C 937 | 14 | 1.4 | 793 | 21 | AAV25517 | Feline Flt-3 ligand |
| C 865 | 14 | 631 | 21 | AAC31898 | Arabidopsis thalia | C 938 | 14 | 1.4 | 793 | 21 | AAC42255 | Arabidopsis thalia |
| C 866 | 14 | 633 | 19 | AAV22068 | DNA encoding a Krt | C 939 | 14 | 1.4 | 795 | 21 | AAV25522 | Feline mature Flt- |
| C 867 | 14 | 633 | 21 | AAA50472 | White shrimp multi | C 940 | 14 | 1.4 | 795 | 21 | AAV25523 | Feline mature Flt- |
| C 868 | 14 | 634 | 21 | AAC48206 | Arabidopsis thalia | C 941 | 14 | 1.4 | 798 | 21 | AAV15953 | Human prostate can |
| C 869 | 14 | 636 | 21 | AAV09598 | Fusarium venenatum | C 942 | 14 | 1.4 | 804 | 21 | AAV01865 | Human colon cancer |
| C 870 | 14 | 636 | 21 | AAV65944 | E. coli proliferat | C 943 | 14 | 1.4 | 804 | 21 | AAV55501 | Canine mature Flt- |
| C 871 | 14 | 637 | 21 | AAV12391 | Aspergillus oryzae | C 944 | 14 | 1.4 | 806 | 21 | AAV55502 | Canine mature Flt- |
| C 872 | 14 | 637 | 21 | AAV12850 | Aspergillus oryzae | C 945 | 14 | 1.4 | 806 | 21 | AAV29177 | Canine mature Flt- |
| C 873 | 14 | 637 | 21 | AAC03783 | Human secreted pro | C 946 | 14 | 1.4 | 810 | 21 | AAC79876 | White pine Pin m I |
| C 874 | 14 | 638 | 21 | AAC60989 | Human vesicle asso | C 947 | 14 | 1.4 | 811 | 8 | AAV70093 | Human secreted pro |
| C 875 | 14 | 640 | 22 | AAV93990 | Primer specific fo | C 948 | 14 | 1.4 | 813 | 20 | AAV17326 | Sequence of the np |
| C 876 | 14 | 641 | 21 | AAV15013 | Trichoderma reesei | C 949 | 14 | 1.4 | 817 | 21 | AAV24967 | Human gene express |
| C 877 | 14 | 642 | 18 | AAV77866 | CDNA encoding huma | C 950 | 14 | 1.4 | 822 | 21 | AAV27326 | Corn beta-carotene |
| C 878 | 14 | 642 | 21 | AAV14799 | Aspergillus oryzae | C 951 | 14 | 1.4 | 825 | 21 | AAV69536 | Human prostate can |
| C 879 | 14 | 643 | 21 | AAV09844 | Fusarium venenatum | C 952 | 14 | 1.4 | 828 | 21 | AAV25505 | Eucalyptus grandis |
| C 880 | 14 | 643 | 21 | AAC47491 | Arabidopsis thalia | C 953 | 14 | 1.4 | 828 | 21 | AAV25506 | Canine Flt-3 ligand |
| C 881 | 14 | 646 | 21 | AAC34315 | Arabidopsis thalia | C 954 | 14 | 1.4 | 830 | 21 | AAC76325 | Human ORFX ORF1880 |
| C 882 | 14 | 649 | 19 | AAV12744 | RBE1 transcription | C 955 | 14 | 1.4 | 841 | 18 | AAV49161 | Human ORFX ORF1880 |
| C 883 | 14 | 655 | 21 | AAC34197 | Arabidopsis thalia | C 956 | 14 | 1.4 | 842 | 22 | AAV59007 | Mouse epididymis-s |
| C 884 | 14 | 655 | 21 | AAC40068 | Arabidopsis thalia | C 957 | 14 | 1.4 | 846 | 21 | AAV236892 | CDNA encoding an a |
| C 885 | 14 | 661 | 20 | AAV88531 | EST clone H0141. | C 958 | 14 | 1.4 | 849 | 21 | AAV62041 | Hydrophobic domain |
| C 886 | 14 | 661 | 21 | AAV9401 | Eucalyptus grandis | C 959 | 14 | 1.4 | 858 | 21 | AAC52817 | Arabidopsis thalia |
| C 887 | 14 | 665 | 21 | AAC79943 | Human secreted pro | C 960 | 14 | 1.4 | 861 | 10 | AAV91840 | Sequence encoding |


```
OY 721 ataaccaacacgctcgggttcctgaagttgctccctcctgacacattcgaaagacattgaa 780
Db 1952 ataaccaacaacgctcgggttcctgaagttgctccctcctgacacattcgaaagacattgaa 2011
OY 781 caactctgaatctgggttttagagacacagatgacagagctgcagcagaacgctccaca 840
Db 2012 caactctgaatctgggttttagagacacagatgacagagctgcagcagaacgctccaca 2071
OY 841 ccgcaaacacagccgagagacacaaagctccggaataatccgtgtttgagacgtacacc 900
Db 2072 ccgcaaacacagccgagagacacaaagctccggaataatccgtgtttgagacgtacacc 2131
OY 901 ctgttctcatctggagatccctctctcgcacacatcagaagggtgtaagcgtgtgtg 960
Db 2132 ctgttctcatctggagatccctctctcgcacacatcagaagggtgtaagcgtgtgtg 2191
OY 961 aacgagacattatgt 976
Db 2192 aacgagacattatgt 2207

RESULT 2
AAT27310
ID AAT27310 standard; cDNA; 3402 BP.
XX
AC AAT27310;
XX
XX 26-NOV-1996 (first entry)
XX
DE Trypanosoma cruzi epimastigotic PTC100t antigen gene.
XX
KW Antigen; Trypanosoma cruzi; epimastigote; serum; Chagas disease; probe;
KM Primer; PCR; polymerase chain reaction; amplification; antibody; ds.
XX
OS Trypanosoma cruzi.
XX
FH Key Location/Qualifiers
FT CDS 266..3013
FT /*CDS= a
FT /product= PTC100t epimastigotic antigen
XX
XX FR2723589-A1.
XX
XX 16-FEB-1996.
XX
XX 12-AUG-1994; 94FR-0010132.
XX
XX 12-AUG-1994; 94FR-0010132.
XX
XX (INMR ) BIO MERIEUX.
XX
XX Jolivet M, Lesenechal M, Paranhos-Baccala G;
XX
XX WPI: 1996-190287/20.
XX
XX P-PSDB; AAR91615.
XX
XX New nucleic acid encoding Trypanosoma cruzi epimastigotic antigen -
XX
XX useful for diagnosis, monitoring and therapy of Chagas disease
XX
XX Claim 1; Page 24-26; 55pp; French.
XX
XX This is the nucleotide sequence encoding a novel isolated antigenic
XX
XX protein from Trypanosoma cruzi epimastigotes, designated PTC100t.
XX
XX The clone Tc50 was isolated from a T. cruzi genomic expression library in
XX
XX lambda gfil, using a mixture of sera from patients with Chagas disease.
XX
XX Clone Tc50 contained an 594 bp insert corresp. to nucleotides 1232-1825
XX
XX of this sequence. The Tc50 sequence was subsequently used to probe a
XX
XX Southern blot of restriction enzyme digested T. cruzi DNA and also screen
XX
XX a lambda gfil library to isolate a 1041 bp EcoRI fragment corresp. to
XX
XX nucleotides 1403-2443 of PTC100t. Primers (AAT27311-5) were synthesised
XX
XX based on the sequences of the 594 and 1041 bp fragments and used to
XX
XX amplify the PTC100t clone as 3 fragments from cDNA derived from mRNA
```

```
CC purified from T. cruzi epimastigotes. The protein or antibodies raised
CC against it can be used in the detection and monitoring of T. cruzi
CC infection i.e., Chagas disease.
XX
XX Sequence 3402 BP; 889 A; 818 C; 958 G; 737 T; 0 other;
SO
Query Match 89.5%; Score 874; DB 17; Length 3402;
Best Local Similarity 99.8%; Pred. NO. 0;
Matches 974; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 caggatcacgctgaacgagcttctgtcctcaatcgttaagccgacgtagctgtcctgtct 60
Db 1232 caggatcacgctgaacgagcttctgtcctcaatcgttaagccgacgtagctgtcctgtct 1291
OY 61 gacatgtcgaatcgtatgtagatcttccatctccggtccctccgcaaggagaacagacag 120
Db 1292 gacatgtcgaatcgtatgtagatcttccatctccggtccctccgcaaggagaacagacag 1351
OY 121 ccagggcccaaaaacatcgtgtgtagtcgacgagcaacccgggtgtgttccctcgggacct 180
Db 1352 ccagggcccaaaaacatcgtgtgtagtcgacgagcaacccgggtgtgttccctcgggacct 1411
OY 181 gacgcgcgagtagagatcacaatacgaactctctccggtcgtcgtacatccctcaca 240
Db 1412 gacgcgcgagtagagatcacaatacgaactctctccggtcgtcgtacatccctcaca 1471
OY 241 cccctgtttcagcgccacgcaagcgacgagccgctctccgcgcgagctgag 300
Db 1472 cccctgtttcagcgccacgcaagcgacgagccgctctccgcgcgagctgag 1531
OY 301 ccgcatcgtggggaggaagatcatctgtaatcagatcagctgggattatgtacc 360
Db 1532 ccgcatcgtggggaggaagatcatctgtaatcagatcagctgggattatgtacc 1591
OY 361 caaaggaagctgtagcaactgagcgccgacgacgaggtctcagcggtgagctcc 420
Db 1592 caaaggaagctgtagcaactgagcgccgacgacgaggtctcagcggtgagctcc 1651
OY 421 acgactacgcccgcagcgaaacaaagctcacaatcagggacaaatgagcagct 480
Db 1652 acgactacgcccgcagcgaaacaaagctcacaatcagggacaaatgagcagct 1711
OY 481 ggaattggtgacgactaattagttgtccagcgcggtcgtctccacagccggtcgaaa 540
Db 1712 ggaattggtgacgactaattagttgtccagcgcggtcgtctccacagccggtcgaaa 1771
OY 541 ccaacagagagaagaagaagcctccgcgcatgtgaacagagctccgttgcgataatgcg 600
Db 1772 ccaacagagagaagaagaagcctccgcgcatgtgaacagagctccgttgcgataatgcg 1831
OY 601 acacgcccgcgtctcaacaacgctctcctccgacgagcgcaacgagatggcttttgcg 660
Db 1832 acacgcccgcgtctcaacaacgctctcctccgacgagcgcaacgagatggcttttgcg 1891
OY 661 gcaagcagtattaccagtcggagggcgaggttcaatcagtcgtgagcggtgagtcgtc 720
Db 1892 gcaagcagtattaccagtcggagggcgaggttcaatcagtcgtgagcggtgagtcgtc 1951
OY 721 ataaccaacacgctcgggttcctgaagttgctccctcctgacacattcgaaagacattgaa 780
Db 1952 ataaccaacaacgctcgggttcctgaagttgctccctcctgacacattcgaaagacattgaa 2011
OY 781 caactctgaatctgggttttagagacacagatgacagagctgcagcagaacgctccaca 840
Db 2012 caactctgaatctgggttttagagacacagatgacagagctgcagcagaacgctccaca 2071
OY 841 ccgcaaacacagccgagagacacaaagctccggaataatccgtgtttgagacgtacacc 900
Db 2072 ccgcaaacacagccgagagacacaaagctccggaataatccgtgtttgagacgtacacc 2131
OY 901 ctgttctcatctggagatccctctcgcacacatcagaagggtgtaagcgtgtgtg 960
```

Db 2132 ctgttctcattgcggaattccctctcgcgaacatcaagaagggtgaagcgttgctg 2191
 QY 961 aacgagccattatgt 976
 |||||||||||||||
 Db 2192 aacgagccattatgt 2207

RESULT 3

AAT27313
 ID AAT27313 standard; cDNA; 22 BP.

XX
 AC AAT27313;

DT 26-NOV-1996 (first entry)

XX T. cruzi epimastigotic Prc100t antigen primer corresp. to bases 1266-87.

XX Antigen: Trypanosoma cruzi; epimastigote; serum; Chagas disease; probe;
 KM Primer; PCR; polymerase chain reaction; amplification; antibody; ss.

XX Synthetic.

OS
 FR2723589-A1.

XX
 PD 16-FEB-1996.

XX 12-AUG-1994; 94FR-0010132.

XX 12-AUG-1994; 94FR-0010132.

XX (INMR) BIO MERIEUX.

XX Jolivet M, Lesenechal M, Paranhos-Baccala G;

XX WPI; 1996-190287/20.

XX PT New nucleic acid encoding Trypanosoma cruzi epimastigotic antigen -
 useful for diagnosis, monitoring and therapy of Chagas disease

XX
 Claim 29; Page 37; 55pp; French.

XX The primers AAT27311-5 were used to PCR amplify the sequence encoding a
 CC novel isolated antigenic protein from Trypanosoma cruzi epimastigotes,
 CC designated Prc100t (AAT27310). The primers, derived from the sequences
 CC of a 594 and 1041 bp fragment of Prc100t, amplified the gene as 3
 CC fragments. This primer corresponds to nucleotides 1266-87 of the Prc100t
 CC sequence, derived from the 594 bp fragment. The 594 bp fragment was
 CC isolated from a T. cruzi genomic expression library in lambda gt11, using
 CC a mixture of sera from patients with Chagas disease. It corresponds to
 CC nucleotides 1232-1825 of Prc100t. The 1041 bp fragment was isolated from
 CC a lambda gt10 library using the 594 bp fragment as a probe. The protein,
 CC or antibodies raised against it, can be used in the detection and
 CC monitoring of T. cruzi infection i.e. Chagas disease.

XX Sequence 22 BP; 3 A; 8 C; 7 G; 4 T; 0 other;

Query Match 2.3%; Score 22; DB 17; Length 22;

Best Local Similarity 100.0%; Pred. No. 0.2; Mismatches 0; Indels 0; Gaps 0;

QY 35 cagccgaggtagctgcgtct 56

Db 1 cagccgaggtagctgcgtct 22

RESULT 4

AAx84095
 ID AAX84095 standard; DNA; 22 BP.

XX
 AC AAX84095;

DT 27-AUG-1999 (first entry)

XX PCR primer for T. cruzi Prc40 coding sequence.
 DE
 XX Prc40; Tc40; infection; diagnosis; immune complex; antigenic determinant;
 KM therapy; antibody; PCR primer; ss.

XX Synthetic.

OS Trypanosoma cruzi.

XX WO9229867-A1.

XX 17-JUN-1999.

XX 10-DEC-1998; 98WO-IB01987.

XX 10-DEC-1997; 97US-0988242.

XX (INMR) BIO MERIEUX.

XX Jolivet M, Lesenechal M, Mandrand B, Paranhos-Baccala G;

XX WPI; 1999-394978/33.

XX New Trypanosoma cruzi antigen

XX Disclosure; Page 21; 65pp; English.

XX This sequence is a PCR primer for DNA encoding the Trypanosoma cruzi
 CC Prc40 protein of the invention. The Prc40 antigenic determinant is
 CC useful as a reagent for detection and/or monitoring of Trypanosoma cruzi
 CC infection from samples including blood serum or plasma, urine, saliva, or
 CC tears, by contacting with the sample and detecting an immune complex. The
 CC Prc40 antigenic determinant, the vector, expression cassette, cell or
 CC antibody are useful for treatment or prevention (vaccine) of a
 CC Trypanosoma cruzi infection in a man or animal. Current Trypanosoma cruzi
 CC antigens are obtained from protein fractions of the noninfectious stage
 CC of the parasite, and these do not allow sufficient production of antigens
 CC for use in reliable serological diagnostic tests. The strain to strain
 CC polymorphism reduces reliability of the tests.

XX Sequence 22 BP; 3 A; 8 C; 7 G; 4 T; 0 other;

Query Match 2.3%; Score 22; DB 20; Length 22;

Best Local Similarity 100.0%; Pred. No. 0.2; Mismatches 0; Indels 0; Gaps 0;

QY 35 cagccgaggtagctgcgtct 56

Db 1 cagccgaggtagctgcgtct 22

RESULT 5

AAT27314/C
 ID AAT27314 standard; cDNA; 21 BP.

XX
 AC AAT27314;

DT 26-NOV-1996 (first entry)

XX T. cruzi epimastigotic Prc100t antigen primer corresp. to bases 2187-2207.

XX Antigen: Trypanosoma cruzi; epimastigote; serum; Chagas disease; probe;

KW Primer; PCR; polymerase chain reaction; amplification; antibody; ss.

XX Synthetic.

OS
 PN FR2723589-A1.

XX 16-FEB-1996.

XX 12-AUG-1994; 94FR-0010132.

PR 12-AUG-1994; 94FR-0010132.
XX
PA (INMR) BIO MERIEUX.
XX
PI Jolivet M, Lesenechal M, Paranhos-Baccala G;
XX
DR WPI; 1996-190287/20.
XX
PT New nucleic acid encoding Trypanosoma cruzi epimastigotic antigen -
PT useful for diagnosis, monitoring and therapy of Chagas disease
XX
PS Claim 29; Page 38; 55pp; French.
XX
CC The primers AAT27311-5 were used to PCR amplify the sequence encoding a
CC novel isolated antigenic protein from Trypanosoma cruzi epimastigotes,
CC designated Prc100t (AAT27310). The primers, derived from the sequences
CC of a 594 and 1041 bp fragment of Prc100t, amplified the gene as 3
CC fragments. This primer corresponds to nucleotides 2187-2207 of the P1100t
CC sequence, derived from the 1041 bp fragment. The 594 bp fragment was
CC isolated from a T. cruzi genomic expression library in lambda g11, using
CC a mixture of sera from patients with Chagas disease. It corresponds to
CC nucleotides 1232-1825 of Prc100t. The 1041 bp fragment was isolated from
CC a lambda g10 library using the 594 bp fragment as a probe. The protein,
CC or antibodies raised against it, can be used in the detection and
CC monitoring of T. cruzi infection i.e. Chagas disease.
XX
SQ Sequence 21 BP; 6 A; 7 C; 3 G; 5 T; 0 other;

Query Match 2.28; Score 21; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.64;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 956 gctgtgaacgagccatcatgt 976
|||||
DB 21 GTGTGAACGAGCCCATATGTC 1

RESULT 6
AAT27315
ID AAT27315 standard; cDNA; 21 BP.
XX
AC AAT27315;
XX
DT 26-NOV-1996 (first entry)
XX
DE T. cruzi epimastigotic Prc100t antigen primer corresp. to bases 1997-2017.
XX
KM Antigen; Trypanosoma cruzi; epimastigote; serum; Chagas disease; probe;
KM Primer; PCR; polymerase chain reaction; amplification; antibody; ss.
XX
OS Synthetic.
XX
XX FR2723589-A1.
XX
XX 16-FEB-1996.
XX
XX 12-AUG-1994; 94FR-0010132.
XX
XX 12-AUG-1994; 94FR-0010132.
XX
XX (INMR) BIO MERIEUX.
XX
XX Jolivet M, Lesenechal M, Paranhos-Baccala G;
XX
XX WPI; 1996-190287/20.
XX
XX New nucleic acid encoding Trypanosoma cruzi epimastigotic antigen -
XX useful for diagnosis, monitoring and therapy of Chagas disease
XX
XX Claim 29; Page 40; 55pp; French.
XX
XX The primers AAT27311-5 were used to PCR amplify the sequence encoding a

CC novel isolated antigenic protein from Trypanosoma cruzi epimastigotes,
CC designated Prc100t (AAT27310). The primers, derived from the sequences
CC of a 594 and 1041 bp fragment of Prc100t, amplified the gene as 3
CC fragments. This primer corresponds to nucleotides 1997-2017 of the P1100t
CC sequence, derived from the 1041 bp fragment. The 594 bp fragment was
CC isolated from a T. cruzi genomic expression library in lambda g11, using
CC a mixture of sera from patients with Chagas disease. It corresponds to
CC nucleotides 1232-1825 of Prc100t. The 1041 bp fragment was isolated from
CC a lambda g10 library using the 594 bp fragment as a probe. The protein,
CC or antibodies raised against it, can be used in the detection and
CC monitoring of T. cruzi infection i.e. Chagas disease.
XX
SQ Sequence 21 BP; 9 A; 5 C; 4 G; 3 T; 0 other;

Query Match 2.28; Score 21; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.64;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 766 cgaagagaccatgaacaact 786
|||||
DB 1 cgaagagaccatgaacaact 21

RESULT 7
AAx84096/c
ID AAx84096 standard; DNA; 21 BP.
XX
AC AAx84096;
XX
DT 27-AUG-1999 (first entry)
XX
DE PCR primer for T. cruzi Prc40 coding sequence.
XX
KM Prc40; Tc40; infection; diagnosis; immune complex; antigenic determinant;
KM therapy; antibody; PCR primer; ss.
XX
OS Synthetic.
XX
XX Trypanosoma cruzi.
XX
XX WO9929867-A1.
XX
XX 17-JUN-1999.
XX
XX 10-DEC-1998; 98WO-IB01987.
XX
XX 10-DEC-1998; 98WO-IB01987.
XX
XX 10-DEC-1997; 97US-0988242.
XX
XX (INMR) BIO MERIEUX.
XX
XX Jolivet M, Lesenechal M, Mandrand B, Paranhos-Baccala G;
XX
XX WPI; 1999-394978/33.
XX
XX New Trypanosoma cruzi antigen
XX
XX Disclosure; Page 21; 65pp; English.
XX
XX This sequence is a PCR primer for DNA encoding the Trypanosoma cruzi
XX Prc40 protein of the invention. The Prc40 antigenic determinant is
XX useful as a reagent for detection and/or monitoring of Trypanosoma cruzi
XX infection from samples including blood serum or plasma, urine, saliva, or
XX tears, by contacting with the sample and detecting an immune complex. The
XX Prc40 antigenic determinant, the vector, expression cassette, cell or
XX antibody are useful for treatment or prevention (vaccine) of a
XX Trypanosoma cruzi infection in a man or animal. Current Trypanosoma cruzi
XX antigens are obtained from protein fractions of the noninfectious stage
XX of the parasite, and these do not allow sufficient production of antigens
XX for use in reliable serological diagnostic tests. The strain to strain
XX polymorphism reduces reliability of the tests.
XX
SQ Sequence 21 BP; 6 A; 7 C; 3 G; 5 T; 0 other;

Query Match 2.2%; Score 21; DB 20; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.64;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 956 ggtgtgaacgaagccattatgt 976
|||||
DB 21 gttctgaacgagccattatgt 1

RESULT 9
AAV30198
ID AAV30198 standard; DNA; 21 BP.
XX
AC AAV30198;
XX
XX
XX 27-AUG-1999 (first entry)
XX
XX PCR primer for T. cruzi PTC40 coding sequence.
XX
XX PTC40; Tc40; infection; diagnosis; immune complex; antigenic determinant;
XX therapy; antibody; PCR primer; ss.
XX
XX Synthetic.
XX Trypanosoma cruzi.
XX
XX WO9229867-A1.
XX
XX 17-JUN-1999.
XX
XX 10-DEC-1998; 98WO-1B01987.
XX
XX 10-DEC-1997; 97US-0988242.
XX
XX (INMR) BIO MERIEUX.
XX
XX Jolivet M, Lesenechal M, Mandrand B, Paranhos-Baccala G;
XX WPI; 1999-394978/33.
XX
XX New Trypanosoma cruzi antigen
XX
XX PS Disclosure; Page 22; 65pp; English.
XX
XX This sequence is a PCR primer for DNA encoding the Trypanosoma cruzi
XX PTC40 protein of the invention. The PTC40 antigenic determinant is
XX useful as a reagent for detection and/or monitoring of Trypanosoma cruzi
XX infection from samples including blood serum or plasma, urine, saliva, or
XX tears, by contacting with the sample and detecting an immune complex. The
XX PTC40 antigenic determinant, the vector, expression cassette, cell or
XX antibody are useful for treatment or prevention (vaccine) of a
XX Trypanosoma cruzi infection in a man or animal. Current Trypanosoma cruzi
XX antigens are obtained from protein fractions of the noninfectious stage
XX of the parasite, and these do not allow sufficient production of antigens
XX for use in reliable serological diagnostic tests. The strain to strain
XX polymorphism reduces reliability of the tests.
XX
XX Sequence 21 BP; 9 A; 5 C; 4 G; 3 T; 0 other;

Query Match 2.2%; Score 21; DB 20; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.64;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 766 cgaagagacatgaacactt 786
|||||
DB 1 cgaagagacatgaacactt 21

RESULT 9
AAV30199/c
ID AAV30199 standard; DNA; 11878 BP.
XX

AAV30199;
XX
XX 14-SEP-1998 (first entry)
XX
XX Protein kinase catalytic subunit gene.
XX
XX Severe combined immunodeficiency disease; SCID; horse; diagnosis;
XX DNA-dependent protein kinase; ds.
XX
XX Equus caballus.
XX
XX WO9821367-A1.
XX
XX 22-MAY-1998.
XX
XX 14-NOV-1997; 97WO-US21066.
XX
XX 15-NOV-1996; 96US-0031261.
XX
XX (TEXA) UNIV TEXAS SYSTEM.
XX
XX Weeks K:
XX
XX WPI; 1998-297967/26.
XX
XX DNA-dependent protein kinase catalytic subunit - useful for
XX determining equine severe combined immunodeficiency alleles
XX
XX PS Disclosure; Page 70-78; 98pp; English.

This isolated DNA molecule encodes an equine DNA-dependent protein
kinase catalytic subunit (DNA-PKcs). A claimed method of
identifying an Arabian horse that is a carrier of equine severe
combined immunodeficiency (SCID) comprises determining whether the
horse has a mutation in a SCID determinant region of the DNA-PKcs
gene (see also AAV30196 and AAV30197). Sequence analysis of DNA-PKcs
genes from normal and SCID equine fibroblasts shows that a 5 bp
deletion is present in SCID foals at a site that corresponds to
nucleotide 9454 of the 12,381 nucleotide coding sequence of the
human transcript. This results in premature termination of the
DNA-PKcs at amino acid 3160 (see AAV56642). Oligonucleotide probes
(see AAV30194 and AAV30195) are provided that precisely span the SCID
determinant region of the DNA-PKcs gene, and which are diagnostic
for the normal and SCID alleles, respectively. Methods are also
provided for identifying for differentiating SCID homozygotes,
heterozygotes and normal horses.

Sequence 11878 BP; 3586 A; 2323 C; 2668 G; 3301 T; 0 other;

Query Match 1.9%; Score 19; DB 19; Length 11878;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 211 actctgcgcgtcgtcat 229
|||||
DB 6446 ACTTCTGCCGCTGTCAT 6428

RESULT 10
AAV30198/c
ID AAV30198 standard; DNA; 11883 BP.
XX
XX AAV30198;
XX
XX 14-SEP-1998 (first entry)
XX
XX Protein kinase catalytic subunit gene.
XX
XX Severe combined immunodeficiency disease; SCID; horse; diagnosis;
XX DNA-dependent protein kinase; ds.
XX
XX Equus caballus.
XX

```

XX      WO9821367-A1.
PN      22-MAY-1998.
XX      14-NOV-1997; 97WO-US21066.
XX      15-NOV-1996; 96US-0031261.
XX      (TEXA ) UNIV TEXAS SYSTEM.
XX      Meeks K;
XX      WPI: 1998-297967/26.
DR      P-PSDB: AAM56642.
XX      DNA-dependent protein kinase catalytic subunit - useful for
PT      determining equine severe combined immunodeficiency alleles
XX      Claim 1; Page 39-44; 98pp; English.
XX      This isolated DNA molecule encodes a DNA-dependent protein kinase
CC      catalytic subunit (DNA-PKcs, see AAM56642) found in Arabian horses.
CC      A claimed method of identifying an Arabian horse that is a carrier
CC      of equine severe combined immunodeficiency (SCID) comprises
CC      determining whether the horse has a mutation in a SCID determinant
CC      region of the DNA-PKcs gene (see also AAV30196 and AAV30197). Sequence
CC      analysis of DNA-PKcs genes from normal and SCID equine fibroblasts
CC      shows that a 5 bp deletion is present in SCID foals at a site that
CC      corresponds to nucleotide 9454 of the 12,381 nucleotide coding
CC      sequence of the human transcript. This results in premature
CC      termination of the DNA-PKsc at amino acid 3160. Oligonucleotide
CC      probes (see AAV30194 and AAV30195) are provided that precisely span the
CC      SCID determinant region of the DNA-PKcs gene, and which are
CC      diagnostic for the normal and SCID alleles, respectively. Methods
CC      are also provided for identifying for differentiating SCID
CC      homozygotes, heterozygotes and normal horses. A claimed plasmid
CC      contains the DNA-PKcs DNA and regulatory elements necessary for
CC      expression of the DNA in a recombinant cell.
XX      Sequence 11883 BP; 3588 A; 2324 C; 2669 G; 3302 T; 0 other;
SQ
Query Match      1.9%; Score 19; DB 19; Length 11883;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      211 actctgcgcgtcgtcat 229
        ||||||||||||||||
DB      6449 ACTTCTGCCGCTGCTGCAT 6431
RESULT 11
AAT27312/C
ID      AAT27312 standard; cDNA; 18 BP.
XX      AAT27312;
AC      AAT27312;
XX      26-NOV-1996 (first entry)
DT      T.cruzi epimastigotic PTC100t antigen primer corresp. to bases 1442-59.
XX      Antigen: Trypanosoma cruzi; epimastigote; serum; Chagas disease; probe;
KW      Primer; PCR; polymerase chain reaction; amplification; antibody; ss.
XX      Synthetic.
OS      Synthetic.
XX      FR2723589-A1.
PN      16-FEB-1996.
XX      12-AUG-1994; 94FR-0010132.
XX

```

```

FR      12-AUG-1994; 94FR-0010132.
XX      (IMMR ) BIO MERIEUX.
XX      Jolivet M, Lesenechal M, Paranhos-Bacala G;
XX      WPI: 1996-190287/20.
XX      New nucleic acid encoding Trypanosoma cruzi epimastigotic antigen -
PT      useful for diagnosis, monitoring and therapy of Chagas disease
XX      Claim 29; Page 36; 55pp; French.
XX      The primers AAT27311-5 were used to PCR amplify the sequence encoding a
CC      novel isolated antigenic protein from Trypanosoma cruzi epimastigotes,
CC      designated PTC100t (AAT27310). The primers, derived from the sequences
CC      of a 594 and 1041 bp fragment of PTC100t, amplified the gene as 3
CC      fragments. This primer corresponds to nucleotides 1442-59 of the PTC100t
CC      sequence, derived from the 594 bp fragment. The 594 bp fragment was
CC      isolated from a T.cruzi genomic expression library in lambda gt11, using
CC      a mixture of sera from patients with Chagas disease. It corresponds to
CC      nucleotides 1232-1825 of PTC100t. The 1041 bp fragment was isolated from
CC      a lambda gt10 library using the 594 bp fragment as a probe. The protein,
CC      or antibodies raised against it, can be used in the detection and
CC      monitoring of T.cruzi infection i.e. Chagas disease.
XX      Sequence 18 BP; 5 A; 4 C; 7 G; 2 T; 0 other;
SQ
Query Match      1.8%; Score 18; DB 17; Length 18;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      211 actctgcgcgtcgtcga 228
        ||||||||||||||||
DB      18 ACTTCTGCCGCTGCTGCA 1
RESULT 12
AAX84094/C
ID      AAX84094 standard; DNA; 18 BP.
XX      AAX84094;
AC      AAX84094;
XX      27-AUG-1999 (first entry)
DT      PCR primer for T. cruzi PTC40 coding sequence.
XX      PTC40; Tc40; infection; diagnosis; immune complex; antigenic determinant;
KW      therapy; antibody; PCR primer; ss.
XX      Synthetic.
OS      Trypanosoma cruzi.
XX      WO9929867-A1.
PN      17-JUN-1999.
XX      10-DEC-1998; 98WO-IB01987.
XX      10-DEC-1997; 97US-0988242.
XX      (IMMR ) BIO MERIEUX.
XX      Jolivet M, Lesenechal M, Mandrand B, Paranhos-Bacala G;
XX      WPI: 1999-394978/33.
XX      New Trypanosoma cruzi antigen
PT      Disclosure; Page 21; 65pp; English.
XX      This sequence is a PCR primer for DNA encoding the Trypanosoma cruzi
CC

```

CC Prt40 protein of the invention. The Prt40 antigenic determinant is
 CC useful as a reagent for detection and/or monitoring of *Trypanosoma cruzi*
 CC infection from samples including blood serum or plasma, urine, saliva, or
 CC tears, by contacting with the sample and detecting an immune complex. The
 CC Prt40 antigenic determinant, the vector, expression cassette, cell or
 CC antibody are useful for treatment or prevention (vaccine) of a
 CC *Trypanosoma cruzi* infection in a man or animal. Current *Trypanosoma cruzi*
 CC antigens are obtained from protein fractions of the noninfectious stage
 CC of the parasite, and these do not allow sufficient production of antigens
 CC for use in reliable serological diagnostic tests. The strain to strain
 CC polymorphism reduces reliability of the tests.

XX Sequence 18 BP: 5 A; 4 C; 7 G; 2 T; 0 other;

Query Match 1.8%; Score 18; DB 20; Length 18;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 211 actctcgcgcgtcgtcga 228
 |||
 Db 18 ACTTCTGCCGCTCTGCA 1

RESULT 13

AAf09309 standard; cDNA: 454 BP.

AAf09309;

13-MAR-2001 (first entry)

Fusarium venenatum EST SEQ ID NO:1832.

Multiple gene expression; filamentous fungal cell; EST;

Expressed sequence tag; *Fusarium venenatum*; *Aspergillus niger*;

Aspergillus oryzae; *Trichoderma reesei*; Identification; recombination;

culture condition; environmental stress; spore morphogenesis;

metabolic pathway engineering; catabolic pathway engineering; ss.

Fusarium venenatum.

WO200056762-A2.

28-SEP-2000.

22-MAR-2000; 2000WO-US07781.

22-MAR-1999; 99US-0273623.

(NOVO) NOVO NORDISK BIOTECH INC.
 (NOVO) NOVO NORDISK AS.

Berta RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB:

WPI: 2000-594572/56.

Monitoring differential expression of genes in filamentous fungal cells.

substrate of expressed sequence tags -

Claim 86; Page 1068; 3161pp; English.

XX The present invention describes a method for monitoring differential
 CC expression of genes in a first filamentous fungal (FF) cell relative to
 CC expression of the same genes in one or more second filamentous fungal
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from
 CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
 CC are used in the methods for monitoring differential expression of genes
 CC in a first filamentous fungal (FF) cell relative to expression of the
 CC same genes in one or more second filamentous fungal cells. Monitoring
 CC the global expression of genes from FF cells allows the production
 CC potential of the microorganisms to be improved. New genes may be

CC discovered, possible functions of unknown open reading frames can be
 CC identified and gene copy number variation and stability can be
 CC monitored. The expression of genes can be used to study how FF cells
 CC adapt to changes in culture conditions, environmental stress, spore
 CC morphogenesis, recombination, metabolic or catabolic pathway
 CC engineering. Using ESTs provides several advantages over genomic or
 CC random cDNA clones including elimination of redundancy as one spot on an
 CC array equals one gene or open reading frame, and organization of the
 CC microarrays based on function of the gene products to facilitate
 CC analysis of the results. AAF07478 to AAF11247 represents ESTs from
 CC *Fusarium venenatum*; AAF11248 to AAF11853 represents ESTs from *Aspergillus*
 CC *niger*; AAF11854 to AAF14878 represents ESTs from *Aspergillus oryzae*; and
 CC AAF14879 to AAF1537 represents ESTs from *Trichoderma reesei*, which are
 CC all specifically claimed in the present invention.

XX Sequence 454 BP: 141 A; 116 C; 89 G; 108 T; 0 other;

Query Match 1.8%; Score 18; DB 21; Length 454;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 220 gctgcgcacccctcga 237
 |||
 Db 205 gctgcgcacccctcga 222

RESULT 14

AAc04783 standard; cDNA: 479 BP.

AAc04783;

06-OCT-2000 (first entry)

Human secreted protein 5' EST, SEQ ID NO: 8858.

Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

gene therapy; chromosome mapping; ss.

Homo sapiens.

EP1033401-A2.

06-SEP-2000.

21-FEB-2000; 2000EP-0200610.

26-FEB-1999; 99US-0122487.

(GESP) GENSET.

Dumas Milne Edwards J, Duclert A, Giordano J:

WPI: 2000-500381/45.

New nucleic acid that is a 5' expressed sequence tag (5' EST) for

obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for

diagnostic, forensic, gene therapy and chromosome mapping procedures -

Claim 1; SEQ ID 8858; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used

CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.

SO Sequence 479 BP; 109 A; 109 C; 99 G; 160 T; 2 other;

Query Match 1.8%; Score 18; DB 21; Length 479;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 492 agcataagtggtgccag 509
|||||
Db 64 agcataagtggtgccag 81

RESULT 15
AAC36458/c
ID AAC36458 standard; DNA; 486 BP.
XX
AC AAC36458;
XX
DE 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 13881.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151338.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 13-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-01593295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161922.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

Query Match 1.8%; Score 18; DB 21; Length 486;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 930 caacatcaggaagggt 947
IIIIIIIIIIIIIIIIII
Db 164 CAACATCAGGAGGGGT 147
RESULT 16
AAC51869/C
ID AAC51869 standard; DNA; 634 BP.
XX
AC AAC51869;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 69916.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 28-APR-1999; 990S-0130891.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 07-MAY-1999; 990S-0132487.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135539.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.

```
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140635.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147302.
PR 04-AUG-1999; 99US-0147304.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161995.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 18; DB 21; Length 634;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 930 Caacatcagaaaggggt 947
DB 165 CAACATCAGAAAGGGGT 148
```

RESULT 17
 ID AAX39943
 AC AAX39943 standard; DNA: 866 BP.
 XX
 AC AAX39943;
 XX
 DT 02-JUL-1999 (first entry)
 XX
 DE Gastric cancer associated gene.
 XX
 KW Cancer associated antigen; diagnosis; research; treatment; human;
 KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
 KW prostate cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9904265-A2.
 XX
 PD 28-JAN-1999.
 XX
 PF 15-JUL-1998; 98WO-US14679.
 XX
 PR 22-JUN-1998; 98US-0102322.
 PR 17-JUL-1997; 97US-0896164.
 PR 10-OCT-1997; 97US-0061599.
 PR 10-OCT-1997; 97US-0061765.
 PR 10-OCT-1997; 97US-0948705.
 PR 11-OCT-1997; 97GB-0021697.
 XX
 PA (LUDM-) LUDMIG INST CANCER RES.
 PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
 PI Pfreundschuh M, Sahlin U, Scanlan MJ, Stockert E;
 PI Tureci O;
 PI WPI: 1999-132448/11.
 XX
 PR New isolated cancer associated nucleic acids and polypeptides -
 PR isolated using sera from cancer patients, used to develop products
 PR for the diagnosis, monitoring or treatment of cancers
 XX
 PS Claim 67; Page 607; 787pp; English.
 XX
 CC The invention relates to a method for diagnosing a disorder characterised
 CC by expression of a human cancer associated antigen precursor coded for by
 CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
 CC biological sample isolated from a subject with an agent that specifically
 CC binds to the NAM, an expression product or a fragment of an expression
 CC product complexed with an HLA molecule; and (b) determining the
 CC interaction between the agent and the NAM or the expression product as a
 CC determination of the disorder. The products and methods can be used in
 CC the diagnosis, monitoring, research, or treatment of conditions
 CC characterised by the expression of various cancer associated antigens.
 CC The invention provides nucleic acid sequences and encoded polypeptides
 CC which are cancer associated antigen precursors expressed in human breast
 CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
 CC lung cancer.
 XX
 SO Sequence 866 BP; 222 A; 235 C; 197 G; 206 T; 6 other;

Query Match 1.8%; Score 18; DB 20; Length 866;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 cccctgcacacccctg 247
 ||||||||||||||||
 DB 37 cccctgcacacccctg 54

RESULT 18

AAC41168/c
 ID AAC41168 standard; DNA: 970 BP.
 XX
 AC AAC41168;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 30900.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129645.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139494.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.

DE Human secreted protein CDNA #44.
 XX Cytostatic; immunosuppressive; nocotropic; neuroprotective; antiviral;
 KM antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
 KM vulnerability; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KM cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KM neurological disease; infection; human; secreted protein; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO200056765-A1.
 XX
 PD 28-SEP-2000.
 XX
 PF 16-MAR-2000; 2000WO-US06823.
 XX
 PR 19-MAR-1999; 99US-0125364.
 XX 08-DEC-1999; 99US-0169623.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Ruben SM, Komatsoulis G;
 XX
 DR MPI: 2000-602215/57.
 DR P-PSDB; AAB34002.
 XX
 PT Nucleic acid molecules encoding human secreted proteins, used in
 PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
 PT Parkinson's diseases and cancers -
 XX
 PS Claim 1: Page 355; 410pp; English.
 XX
 SQ The invention relates to the isolation of genes AAC59392-C59439 encoding
 CC 48 human secreted proteins AAB3963-B34006. The genes can be used to
 CC generate fusion proteins by linking to the gene for the human
 CC immunoglobulin G Fc portion (SEQID1) for increasing the stability of
 CC the fusion protein as compared to the human protein only. The genes and
 CC proteins are useful for preventing, ameliorating or treating medical
 CC conditions, e.g. by protein or gene therapy. The genes are isolated
 CC from a range of human tissues disclosed in the specification. The
 CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer, and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
 CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC
 SQ Sequence 1302 BP; 249 A; 470 C; 338 G; 230 T; 15 other;

Query Match 1.8%; Score 18; DB 21; Length 1302;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 44 gttagctgcctcgtgctg 61
 |||
 Db 669 gttagctgcctcgtgctg 652

RESULT 20
 ID AAQ77884
 XX AAQ77884 standard; cDNA: 2520 BP.
 AC AAQ77884;
 XX
 XX 06-JUL-1995 (first entry)
 DT
 XX Neural thread protein AD16c CDNA.
 DE

XX
 KM Neural thread protein AD16c; Alzheimer's; neuroectodermal tumours;
 KM malignant astrocytomas; glioblastomas; ss.
 XX
 OS Rattus rattus.
 XX
 PN W09423756-A.
 XX
 PD 27-OCT-1994.
 XX
 PF 20-APR-1994; 94WO-US04321.
 XX
 PR 20-APR-1993; 93US-0050559.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI De la Monte SM, Wands JR;
 XX
 DR MPI: 1994-341497/42.
 XX
 PT Detection of neural thread proteins - to detect sporadic and
 PT familial Alzheimer's disease, neuroectodermal tumours, malignant
 PT astrocytomas and glioblastomas (Eng).
 XX
 PS Example 4; Fig 165; 158pp; English.
 XX
 SQ AAQ77884 is the AD16c neural thread protein (NTP) cDNA. This
 CC sequence was used in the development of an antibody dependent
 CC method, for the detection of NTPs. This new method could be
 CC used to diagnose Alzheimer's disease (differentiating between
 CC sporadic and familial), neuroectodermal tumours, malignant
 CC astrocytomas and glioblastomas.
 CC
 SQ Sequence 2520 BP; 661 A; 577 C; 581 G; 701 T; 0 other;

Query Match 1.8%; Score 18; DB 15; Length 2520;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 230 cccctgcatacccccctg 247
 |||
 Db 298 cccctgcatacccccctg 315

RESULT 21
 ID AAT27766
 XX AAT27766 standard; cDNA: 2520 BP.
 AC AAT27766;
 XX
 XX 14-NOV-1996 (first entry)
 DT
 XX AD 16c human neural thread protein clone (partial sequence).
 DE
 KM Neural thread protein; NTP; diagnosis; detection;
 KM Alzheimer's disease; neuroectodermal tumour; malignant astrocytoma;
 KM monoclonal antibody; binding fragment; ds.
 XX
 OS Homo sapiens.
 XX
 PN W09615272-A1.
 XX
 PD 23-MAY-1996.
 XX
 PF 14-NOV-1995; 95WO-US17111.
 XX
 PR 14-NOV-1994; 94US-0340426.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI De la Monte S, Wands JR;
 XX

| | |
|---|---|
| DR | WP1: 1996-259865/26. |
| XX | Detection of neural thread protein in diagnosis of Alzheimer's |
| PT | disease - also NTP DNA and protein sequences used in gene and |
| PT | anti:sense therapy |
| XX | |
| PS | Example 4c; Figure 16S; 238pp; English. |
| CC | A method for detecting the presence of neural thread protein (NTP) |
| CC | having a molecular weight of 8, 14, 17, 21, 26 or 42 kD in a human |
| CC | subject comprises (a) contacting a sample from a human subject that |
| CC | is suspected of containing the NTP with at least one molecule |
| CC | capable of binding to the protein; and (b) detecting any of the |
| CC | molecule bound to the protein. The binding molecule is selected |
| CC | from an antibody free of natural impurities, a monoclonal antibody |
| CC | or a binding fragment of either of these. The method may be used for |
| CC | diagnosing the presence of Alzheimer's disease, neuroectodermal |
| CC | tumours and a malignant astrocytoma in a human. A number of clones |
| CC | of neural thread protein were isolated from healthy 17-18 week old |
| CC | fœtal human brain (HB) 2 year old temporal lobe neocortex and end |
| CC | stage Alzheimer's disease (AD) cerebral cortex. See AA72753-75. |
| SQ | |
| SQ | Sequence 2520 BP; 661 A; 577 C; 581 G; 701 T; 0 other; |
| OY | Query Match 1.88; Score 18; DB 17; Length 2520; |
| Bst Local Similarity 100.0%; Pred. No. 19; | |
| Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| OY | 230 cccctgcacccccctg 247 |
| Db | 298 cccctgcacccccctg 315 |
| RESULT 22 | |
| AAZ55887/c | |
| ID | AAZ55887 standard; DNA; 68750 BP. |
| XX | |
| AC | AAZ55887; |
| XX | |
| DT | 10-APR-2000 (first entry) |
| XX | |
| DE | Sorangium cellulosum 68.75 kb contig. |
| XX | |
| KW | Epithallone biosynthesis; type I polyketide synthase; taxol substitute; |
| RW | anticancer; ds. |
| XX | |
| OS | Sorangium cellulosum. |
| XX | |
| FH | Key Location/Qualifiers |
| FT | CDS 1..1826 |
| FT | /tag= a |
| FT | /partial |
| FT | /product= "Partial Orf 1 protein (AAZ558580)" |
| FT | /note= "No initiation codon given in the specification" |
| FT | complement (1900..3171) |
| FT | /tag= b |
| FT | /product= "Orf 2 protein (AAZ558581)" |
| FT | 3415..5556 |
| FT | /tag= c |
| FT | /product= "Orf 3 protein (AAZ558582)" |
| FT | complement (5612..5992) |
| FT | /tag= d |
| FT | /product= "Orf 4 protein (AAZ558583)" |
| FT | 6226..6675 |
| FT | /tag= e |
| FT | /product= "Orf 5 protein (AAZ558584)" |
| FT | 7610..11875 |
| FT | /tag= f |
| FT | /product= "Type I polyketide synthase, EPOS A |
| FT | (AAZ558573)" CDS 11872..116104 |
| FT | /tag= g |
| FT | /product= "Non-ribosomal peptide synthetase, EPOS P |

| | | |
|----|--|--|
| FT | | (AAV58574) " |
| FT | CDS | 16251..21749 |
| FT | | /tag= h |
| FT | | /product= "Type I polyketide synthase, EPOS B |
| FT | (AAV58575) " | CDS 21746..43519 |
| FT | | /tag= i |
| FT | | /product= "Type I polyketide synthase, EPOS C |
| FT | (AAV58576) " | CDS 43524..54920 |
| FT | | /tag= j |
| FT | | /product= "Type I polyketide synthase, EPOS D |
| FT | (AAV58577) " | CDS 54935..62254 |
| FT | | /tag= k |
| FT | | /product= "Type I polyketide synthase, EPOS E |
| FT | (AAV58578) " | CDS 62369..63628 |
| FT | | /tag= l |
| FT | | /product= "Cytochrome P450 oxygenase homologue, EPOS F |
| FT | | (AAV58579) " |
| FT | CDS | 63779..64333 |
| FT | | /tag= m |
| FT | | /product= "Orf 6 protein (AAV58585) " |
| FT | CDS | complement (63853..64290) |
| FT | | /tag= n |
| FT | | /product= "Orf 7 protein (AAV58586) " |
| FT | CDS | 64363..64920 |
| FT | | /tag= o |
| FT | | /product= "Orf 8 protein (AAV58587) " |
| FT | CDS | complement (64287..64727) |
| FT | | /tag= p |
| FT | | /product= "Orf 9 protein (AAV58588) " |
| FT | CDS | 65063..65767 |
| FT | | /tag= q |
| FT | | /product= "Orf 10 protein" |
| FT | CDS | complement (65008..65874) |
| FT | | /tag= r |
| FT | | /product= "Orf 11 protein (AAV58590) " |
| FT | CDS | complement (65871..66338) |
| FT | | /tag= s |
| FT | | /product= "Orf 12 protein (AAV58591) " |
| FT | CDS | 66667..67137 |
| FT | | /tag= t |
| FT | | /product= "Orf 13 protein (AAV58592) " |
| FT | CDS | 67334..68251 |
| FT | | /tag= u |
| FT | | /product= "Orf 14 protein (AAV58593) " |
| FT | CDS | 68346..68750 |
| FT | | /partial |
| FT | | /tag= v |
| FT | | /product= "Partial Orf 15 protein (AAV58594) " |
| FT | | /note= "No termination codon given in the specification" |
| XX | | |
| PN | W09966028-A2. | |
| PD | 23-DEC-1999. | |
| PF | 16-JUN-1999; | 99WO-EP04171. |
| PR | 18-JUN-1998; | 98US-0099504. |
| PR | 24-SEP-1998; | 98US-0101631. |
| PR | 05-FEB-1999; | 99US-0118906. |
| XX | | |
| PA | (NOVS) NOVARTIS AG. | |
| PA | (NOVS) NOVARTIS-ERFINDUNGEN VERN GES MBH. | |
| PI | Schupp T, Ligon JM, Molnar I, Zirkle R, Goerlach J, Cyr D; | |
| XX | | |
| DR | WPI: 2000-097741/08. | |
| DR | P-PSDB; AAV58573, AAY58574, AAY58575, AAY58576, AAY58577, AAY58578, | |
| DR | AAV58579, AAY58580, AAY58581, AAY58582, AAY58583, AAY58584, | |
| DR | AAV58585, AAY58586, AAY58587, AAY58588, AAY58590, AAY58591, | |
| DR | AAV58592, AAY58593, AAY58594. | |
| XX | | |
| PT | New isolated epoithione synthase genes, used for the recombinant production of epoithione for use in cancer therapy - | |

XX Claim 14; Page 87-104; 174pp; English.
PS This sequence represents a 68.75 kb contig from *Sorangium cellulosum*
XX comprising 22 open reading frames (ORFs) and includes genes encoding
CC proteins involved in the biosynthesis of epothilones. Epothilones A and
CC B are 16-membered macrocyclic polyketides with an acylcysteine-derived
CC starter unit; polyketides being synthesised from two-carbon building
CC blocks, the beta-carbon of which always carries a keto group. Each round
CC of two-carbon addition is carried out by a complex of enzymes known as
CC the polyketide synthase in a manner similar to fatty acid biosynthesis.
CC EPoS A (AAV58573) and EPoS P (AAV58574) are involved in formation of
CC the thiazole ring formation of epothilones, and EPoS B, EPoS C, EPoS D
CC and EPoS E (AAV58575-V58578) are involved in polyketide backbone
CC formation. EPoS F (AAV58579) is an epothilone macrolactone oxidase, and
CC the proteins Orf 3 (AAV58582) and Orf14 (AAV58593) are thought to be
CC involved in transport. Epothilones mimic the biological activity of
CC taxol, and may be substituted for taxol in cancer chemotherapeutic
CC compositions. Epothilones exhibit a much lower drop in potency against a
CC multiply drug-resistant cell line compared with taxol, and are
CC considerably less efficiently exported from such cells by the multidrug
CC resistance protein (MDR, or P-glycoprotein). Despite the potential of
CC epothilones as anticancer agents, they are problematical to produce on a
CC large scale. Epothilones are too complex for industrial scale chemical
CC synthesis, and *Sorangium cellulosum* is difficult to ferment, producing
CC poor yields of epothilones. The nucleic acids of the invention may be
CC used for the recombinant production of epothilones in a heterologous host
CC that is more amenable to fermentation.
XX
SQ Sequence 66750 BP; 9596 A; 22458 C; 25537 G; 11159 T; 0 other;

Query Match 1.8%; Score 18; DB 21; Length 68750;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 517 tcgtctccacagcgcg 534
IIIIIIIIIIIIIIIIIIII
DB 19020 TCCTCTCCACACGCCGC 19003

RESULT 23
AAA29349/c
ID AAA29349 standard; DNA; 71989 BP.
XX
AC AAA29349;
XX
DT 12-SEP-2000 (first entry)
XX
DE *Sorangium cellulosum* epothilone polyketide synthase operon genomic DNA.
XX
XX Epothilone; polyketide synthase; epOA; epOB; epOC; epOD; epOE; epOF;
KW epOL; epOK; P450 epoxidase; ORFA; ORFB; promoter; enhancer; anti-fungal;
KM tubulin polymerization assay; anti-tumour; cytostatic; ds.
XX
XX *Sorangium cellulosum*.
OS
XX
FH Key Location/Qualifiers
FT 3..992
FT /tag= a
FT /label= ORF A
FT /product= transposase
FT /note= "not part of the PKS"
FT 989..1501
FT /tag= b
FT /label= ORF B
FT /product= transposase
FT /note= "not part of the PKS"
FT 1398..6263
FT /tag= c
FT /label= epOA_gene
FT /note= "encodes the loading domain"
FT 2031..3548
FT misc_RNA

FT /tag= d
FT /note= "encodes ketide synthase (KS-Y) of the loading
FT domain"
FT 3621..4661
FT /tag= e
FT /note= "encodes acyl transferase (AT) of the loading
FT domain"
FT 4917..5810
FT /tag= f
FT /note= "encodes enoyl reductase (ER) of the loading
FT domain, potentially involved in formation of the
FT thiazole moiety"
FT 5856..6135
FT /tag= g
FT /note= "encodes acyl carrier protein (ACP) of the loading
FT domain"
FT 6260..10493
FT /tag= h
FT /label= epOB_gene
FT /note= "encodes module 1, the NRPS module"
FT 2031..3548
FT /tag= i
FT /note= "encodes condensation domain C2 of the NRPS
FT module"
FT 2031..3548
FT /tag= j
FT /note= "encodes condensation domain C2 of the NRPS
FT module"
FT 6861..6887
FT /tag= k
FT /note= "encodes heterocyclization signature sequence"
FT 6861..6887
FT /tag= l
FT /note= "encodes condensation domain C4 of the NRPS
FT module"
FT 7358..7366
FT /tag= m
FT /note= "encodes condensation domain C7 (partial) of the
FT NRPS module"
FT 7898..7921
FT /tag= n
FT /note= "encodes adenylation domain A1 of the NRPS module"
FT 7898..7921
FT /tag= o
FT /note= "encodes adenylation domain A1 of the NRPS module"
FT 8261..8308
FT /tag= p
FT /note= "encodes adenylation domain A3 of the NRPS module"
FT 8411..8422
FT /tag= q
FT /note= "encodes adenylation domain A4 of the NRPS module"
FT 8861..8905
FT /tag= r
FT /note= "encodes adenylation domain A6 of the NRPS module"
FT 8966..8983
FT /tag= s
FT /note= "encodes adenylation domain A7 of the NRPS module"
FT 9090..9179
FT /tag= t
FT /note= "encodes adenylation domain A8 of the NRPS module"
FT 9183..9992
FT /tag= u
FT /note= "encodes oxidation region for forming thiazole"
FT 10121..10138
FT /tag= v
FT /note= "encodes adenylation domain A10 of the NRPS
FT module"
FT 10261..10306
FT /tag= w
FT /note= "encodes thiolation domain (PCP) of the NRPS
FT module"
FT 10639..16137
FT /tag= x


```

FT      /label= epoc_gene
FT      /note= "encodes module 2"
FT      10654..12033
FT      /*tag= y
FT      /note= "encodes KS2, the KS domain of module 2"
FT      12250..13287
FT      /*tag= z
FT      /note= "encodes AT2, the AT domain of module 2"
FT      13327..13899
FT      /*tag= aa
FT      /note= "encodes dehydratase (DH) 2, the DH domain of
FT      module 2"
FT      14962..15756
FT      /*tag= ab
FT      /note= "encodes ketoreductase (KR) 2, the KR domain of
FT      module 2"
FT      15763..16008
FT      /*tag= ac
FT      /note= "encodes ACP2, the ACP domain of module 2"
FT      16134..37907
FT      /*tag= ad
FT      /label= epod_gene
FT      /note= "encodes modules 3-6"
FT      16425..17606
FT      /*tag= ae
FT      /note= "encodes KS3"
FT      17817..18857
FT      /*tag= af
FT      /note= "encodes AT3"
FT      19581..20396
FT      /*tag= ag
FT      /note= "encodes KR3"
FT      20424..20642
FT      /*tag= ah
FT      /note= "encodes ACP3"
FT      20706..22082
FT      /*tag= ai
FT      /note= "encodes KS4"
FT      22296..23336
FT      /*tag= aj
FT      /note= "encodes AT4"
FT      24069..24647
FT      /*tag= ak
FT      /note= "encodes KR4"
FT      24867..25151
FT      /*tag= al
FT      /note= "encodes ACP4"
FT      25203..26576
FT      /*tag= am
FT      /note= "encodes KS5"
FT      26793..27883
FT      /*tag= an
FT      /note= "encodes AT5"
FT      27966..28574
FT      /*tag= ao
FT      /note= "encodes DH5"
FT      29433..30287
FT      /*tag= ap
FT      /note= "encodes ER5"
FT      30321..30869
FT      /*tag= aq
FT      /note= "encodes KR5"
FT      31077..31373
FT      /*tag= ar
FT      /note= "encodes ACP5"
FT      31440..32807
FT      /*tag= as
FT      /note= "encodes KS6"
FT      33018..34067
FT      /*tag= at
FT      /note= "encodes AT6"
FT      34107..34676
FT      /*tag= au

```

```

FT      /note= "encodes DH6"
FT      35760..36641
FT      /*tag= av
FT      /note= "encodes ER6"
FT      36705..37256
FT      /*tag= aw
FT      /note= "encodes KR6"
FT      37470..37769
FT      /*tag= ax
FT      /note= "encodes ACP6"
FT      37912..49308
FT      /*tag= ay
FT      /label= epoe_gene
FT      /note= "encodes modules 7 and 8"
FT      38014..39375
FT      /*tag= az
FT      /note= "encodes KS7"
FT      39589..40626
FT      /*tag= ba
FT      /note= "encodes AT7"
FT      41341..41922
FT      /*tag= bb
FT      /note= "encodes KR7"
FT      42181..42423
FT      misc_RNA

Query Match      1.88; Score 18; DB 21: Length 71989;
Best Local Similarity 100.08; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      517  tgcgtctccacagccgcg 534
          |||f|||||
Db      13408  TCGTCTCCACAGCCGCG 13391

RESULT  24
ID      AAX84320
AC      AAX84320;
XX      08-SEP-1999 (first entry)
XX      DE      Stealth virus nucleic acid clone, SPQ ID NO: 12.
XX      KW      Stealth virus; detection; diagnosis; infection; ss.
XX      OS      Stealth virus.
XX      PN      WO9934019-A1.
XX      PD      08-JUL-1999.
XX      PE      30-DEC-1998; 98WO-US27744.
XX      PR      30-DEC-1997; 97US-0001184.
XX      PA      (MART/) MARTIN W J.
XX      PI      Martin WJ;
XX      DR      WPI; 1999-405521/34.
XX      PS      Novel strains of stealth virus
XX      PS      Claim 19; Page 44; 95pp; English.
CC      This sequence represents a Stealth virus nucleic acid clone. The
CC      invention relates to a method of detecting and characterizing a stealth
CC      virus by reacting a sample suspected of containing a stealth virus with a
CC      probe from a known stealth virus and sequencing the resultant isolated
CC      nucleotide. The method comprises the steps of: (a) isolating DNA or RNA
CC      from a sample suspected of containing a stealth virus, e.g. a culture of
CC      cells showing a viral cytopathic effect; (b) testing the reactivity of

```

CC the isolated DNA or RNA with a molecular probe that contains at least 18
CC or more contiguous nucleotides identical to sequence previously
CC identified from a stealth virus; and, optionally (c) sequencing the
CC isolated DNA or RNA molecules that react with the probe. The method is
CC used to detect stealth virus in a biological product, food or in the
CC environment. The method is also used to evaluate agents for their
CC inhibitory or stimulatory effects on stealth virus replication and to
CC determine capacity of the virus to recombine with and potentially alter
CC the nucleic acid sequences of a cell or bacterium.
XX
SQ Sequence 198 BP; 40 A; 71 C; 54 G; 33 T; 0 other;

Query Match 1.7%; Score 17; DB 20; Length 198;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 395 cgacgaggtctacggcg 411
|||||
DB 177 cgacgaggtctacggcg 193

RESULT 25

AAQ26088/c
ID AAQ26088 standard; DNA; 334 BP.

XX
AC AAQ26088;

XX
DT 24-DEC-1992 (first entry)

XX
DE MIG-2.

XX
KM Intron-free DNA; monocytic induced by interferon-gamma; cytokine;

XX
KW IFN-gamma; ds.

XX
OS Mammalian.

XX
FH Key Location/Qualifiers
40..334

XX
FT CDS /tag- a

XX
PN MO9210582-A.

XX
PD 25-JUN-1992.

XX
PE 09-DEC-1991; 91MO-US09143.

XX
PR 10-DEC-1990; 90US-0624742.

XX
PA (FARB/) FARBER J M.

XX
PI Farber JM;

XX
DR WPI: 1992-234637/28.

XX
DR P-PSDB; AAR25341.

XX
PT New monoclonal induced by IFN-gamma (MIGs) - involved in response
PT to immunological and inflammatory stimuli. probes contg. DNA
PT complementary to DNA encoding MIG and antibodies to MIG are used
PT to assay IFN-gamma

XX
PS Disclosure; Fig 5; 51pp; English.

XX
CC The sequence given is an intron-free DNA encoding mammalian MIG-II (
CC monocytic induced by interferon-gamma). This protein encoded by this
CC sequence hybridizes under conditions of high stringency to MIG-I (
CC See also AAQ26087). MIG-II is a new mammalian cytokine which is a
CC member of the platelet factor 4 family. It has potential
CC therapeutic value as it can be used to mimic some IFN-gamma
CC activities.
XX
SQ Sequence 334 BP; 110 A; 74 C; 72 G; 78 T; 0 other;

Query Match 1.7%; Score 17; DB 13; Length 334;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 787 ctgaatcgtggttaga 803
|||||
DB 277 CTGAATCTGGTTTGA 261

RESULT 26

AAC09227/c
ID AAC09227 standard; cDNA; 353 BP.

XX
AC AAC09227;

XX
DT 06-OCT-2000 (first entry)

XX
DE Human secreted protein 5' EST, SEQ ID NO: 13302.

XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

XX
OS Homo sapiens.

XX
PN EP1033401-A2.

XX
PD 06-SEP-2000.

XX
PF 21-FEB-2000; 2000EP-0200610.

XX
PR 26-FEB-1999; 99US-0122487.

XX
PA (GEST) GENSET.

XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX
DR WPI: 2000-500381/45.

XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 13302; 71pp + CD-ROM; English.

XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 353 BP; 94 A; 74 C; 57 G; 127 T; 1 other;

Query Match 1.7%; Score 17; DB 21; Length 353;
Best Local Similarity 100.0%; Pred. No. 66;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 352 aatgtaccaccaaggag 368
|||||
DB 68 AATGTACCCCAAGAG 52

RESULT 27

AAC00184/C
ID AAC00184 standard; cDNA: 364 BP.
XX
AC AAC00184;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 182.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KM gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EPI033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI: 2000-500381/45.
DR P-PSDB; AAG00178.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 182; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 364 BP; 125 A; 78 C; 74 G; 83 T; 4 other;

Query Match 1.7%; Score 17; DB 21; Length 364;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 787 ctgaacttggttaga 803
|||
DB 300 CTGAATCTGGGTTTGA 284

RESULT 28
AAC01025
ID AAC01025 standard; cDNA: 386 BP.
XX
AC AAC01025;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 1023.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EPI033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI: 2000-500381/45.
DR P-PSDB; AAG01019.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 1023; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 386 BP; 125 A; 88 C; 102 G; 70 T; 1 other;

Query Match 1.7%; Score 17; DB 21; Length 386;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 aacagcagcgagccaa 129
|||
DB 92 aacgagcagcgagccaa 108

RESULT 29
AAC38084
ID AAC38084 standard; DNA: 513 BP.
XX
AC AAC38084;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 19733.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 28-APR-1999; 990S-0130891.
PR 30-APR-1999; 990S-0131449.
PR 04-MAY-1999; 990S-0132048.
PR 04-MAY-1999; 990S-01322407.
PR 05-MAY-1999; 990S-0132484.
PR 06-MAY-1999; 990S-0132486.
PR 07-MAY-1999; 990S-0132487.
PR 11-MAY-1999; 990S-0132863.
PR 14-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137502.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.

PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144684.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 03-AUG-1999; 990S-0146389.
PR 04-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-015138.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.

29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.7%; Score 17; DB 21; Length 513;

Best Local Similarity 100.0%; Pred. No. 65;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 222 tgctgcacccctgcac 238
|||||

Db 481 tgctgcacccctgcac 497

RESULT 30

AAC93440/c
ID AAC93440 standard; cDNA: 552 BP.

XX AAC93440;

DT 16-FEB-2001 (first entry)

XX Human secreted protein gene 19 SEQ ID NO:29.

XX Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;
XX antiproliferative; cytoskeletal; cardiant; vasotropic; cerebroprotective;
XX neurotropic; neuroprotective; antibacterial; virucide; fungicide;
XX ophthalmological; vulnery; autoimmune disease; Rheumatoid arthritis;
XX hyperproliferative disorders; cancer; cardiovascular disorder;
XX cardiac arrest; cerebrovascular disorder; nervous system disorder;
XX Alzheimer's disease; ocular disorder; wound healing; skin aging; ss.

XX Homo sapiens.

XX WO200061625-A1.

XX 19-OCT-2000.

XX 06-APR-2000; 2000MO-US08981.

XX 09-APR-1999; 99US-0128701.

XX 20-JAN-2000; 2000US-0177166.

XX (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX Rosen CA, Ruben SM, Komatsoulis G;
PI WPI; 2000-619226/59.
DR P-PSDB; AAB51747.
XX New nucleic acid molecules encoding 48 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -
XX
PS Claim 1; Page 423-424; 500pp; English.

XX Polynucleotide sequences AAC93442 - AAC93449 represent cDNA encoding
CC human secreted proteins AAB51724 - AAB51777. Sequences AAB51778 -
CC AAB51825 represent alternative polypeptides encoded by the genes, and
CC amino acid sequences to which they are homologous. The genes and proteins
CC have activities dependent on the tissues and cells in which they are
CC expressed. Examples of their activities include immunosuppressive;
CC antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;
CC vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;
CC virucide; fungicide; ophthalmological; and vulnery. The secreted
CC proteins, polynucleotides, antagonists and agonists may be useful in
CC treating, preventing and/or diagnosing diseases and disorders such as
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,
CC infections caused by bacteria, viruses and fungi and ocular disorders
CC e.g. corneal infection. The polypeptides can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,
CC carbohydrate, vitamins, minerals, cofactors and other nutritional
CC components. Oligonucleotide AAC93413 - AAC93421 and peptide AAB51723 are
CC used in the isolation and characterization of the proteins and
XX polynucleotides of the invention.

SQ Sequence 552 BP; 94 A;167 C; 181 G; 102 T; 8 other;

Query Match 1.7%; Score 17; DB 21; Length 552;

Best Local Similarity 100.0%; Pred. No. 65;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 680 agggcgaggttcacag 696
|||||

Db 356 AGGCGAGGTTCACAG 340

RESULT 31

AAF13645/c
ID AAF13645 standard; cDNA: 662 BP.

XX AAF13645;

DT 13-MAR-2001 (first entry)

XX Aspergillus oryzae EST SEQ ID NO:6168.

XX Multiple gene expression; filamentous fungal cell; EST;

XX expressed sequence tag; Fusarium venenatum; Aspergillus niger;

XX Aspergillus oryzae; Trichoderma reesei; identification; recombination;

XX culture condition; environmental stress; spore morphogenesis;

XX metabolic pathway engineering; catabolic pathway engineering; ss.

XX Aspergillus oryzae.

XX WO200056762-A2.

XX 28-SEP-2000.
 XX 22-MAR-2000; 2000WO-US07781.
 XX 22-MAR-1999; 99US-0273623.
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 XX (NOVO) NOVO NORDISK AS.
 XX Berka RM, Rey MM, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
 XX WPI: 2000-594572/56.
 XX Monitoring differential expression of genes in filamentous fungal cells
 XX uses fluorescence-labeled nucleic acids isolated from the cells and a
 XX substrate of expressed sequence tags -
 XX Claim 88; Page 2541; 3161pp; English.
 XX The present invention describes a method for monitoring differential
 XX expression of genes in a first filamentous fungal (FF) cell relative to
 XX expression of the same genes in one or more second filamentous fungal
 XX cells. The method uses fluorescence-labeled nucleic acids isolated from
 XX the FF cells and a substrate of expressed sequence tags (EST). The ESTs
 XX are used in the methods for monitoring differential expression of genes
 XX in a first filamentous fungal (FF) cell relative to expression of the
 XX same genes in one or more second filamentous fungal cells. Monitoring
 XX the global expression of genes from FF cells allows the production
 XX potential of the microorganisms to be improved. New genes may be
 XX discovered, possible functions of unknown open reading frames can be
 XX identified and gene copy number variation and stability can be
 XX monitored. The expression of genes can be used to study how FF cells
 XX adapt to changes in culture conditions, environmental stress, spore
 XX morphogenesis, recombination, metabolic or catabolic pathway
 XX engineering. Using ESTs provides several advantages over genomic or
 XX random cDNA clones including elimination of redundancy as one spot on an
 XX array equals one gene or open reading frame, and organisation of the
 XX microarrays based on function of the gene products to facilitate
 XX analysis of the results. AAF07478 to AAF11247 represents ESTs from
 XX Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus
 XX niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and
 XX AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are
 XX all specifically claimed in the present invention.
 XX Sequence 662 BP; 148 A; 168 C; 175 G; 171 T; 0 other;
 XX SQ

Query Match 1.7%; Score 17; DB 21; Length 662;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 712 gattcgcataaccaa 728
 ||||||||||||||||
 DB 378 GAGTCGTCATACCAA 362

RESULT 32
 AAV16884
 ID AAV16884 standard; DNA; 673 BP.
 XX AAV16884;
 AC
 XX 07-AUG-1998 (first entry)
 XX
 DE Human prostate cancer marker UC Band #31.
 XX
 KW Prostate cancer; human; marker; diagnosis; treatment; probe; ss.
 XX Homo sapiens.
 OS
 XX
 PN W09804689-A1.
 XX

PD 05-FEB-1998.
 XX 31-JUL-1996; 96WO-US12516.
 XX 31-JUL-1996; 96WO-US12516.
 XX (UROC-) UROCOR INC.
 XX An G, O'hara SM, Ralph D, Veltri R;
 XX WPI: 1998-130681/12.
 XX Human prostate cancer marker - useful for detection and treatment of
 XX human prostate cancer
 XX Claim 1; Page 135; 229pp; English.
 XX This represents a marker sequence for human prostate cancer. Isolated
 XX nucleic acid segments shown in AAV16881 to AAV16885, AAV16890 to
 XX AAV16903, AAV26351 and AAV26352 which can act as human prostate cancer
 XX markers are provided in the specification. It also provides methods for
 XX identifying markers for human prostate cancer and for detection of
 XX prostate cancer cells. The markers can be identified by amplifying human
 XX prostate RNA to provide nucleic acid amplification products, separating
 XX the products and identifying those RNA that are differentially expressed
 XX between human prostate cancers versus normal or benign human prostate.
 XX Prostate cancer cells in a sample can be detected by detecting a nucleic
 XX acid in a sample, the nucleic acid being a prostate cancer marker.
 XX Primers and probes derived from this marker can be used for the detection
 XX of prostate cancer cells in a sample. Antibodies against the protein
 XX encoded by the marker nucleic acid fragments, inhibitors of the protein
 XX and oligonucleotides antisense to the markers can be used in the
 XX treatment of prostate cancer. The antibodies can also be used for the
 XX diagnosis of human prostate cancer.
 XX Sequence 673 BP; 222 A; 140 C; 145 G; 166 T; 0 other;
 XX SQ

Query Match 1.7%; Score 17; DB 19; Length 673;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 537 gaaccaacagagaag 553
 ||||||||||||||||
 DB 542 gaaccaacagagaag 558

RESULT 33
 AAX26019
 ID AAX26019 standard; DNA; 673 BP.
 XX AAX26019;
 AC
 XX 20-MAY-1999 (first entry)
 XX
 DE Prostate disease marker gene fragment UC Band #31.
 XX
 KW Prostate cancer; benign prostatic hyperplasia; marker gene; tumour;
 XX differentiation; Reverse Transcription Polymerase Chain Reaction;
 XX diagnostic; progression; cancer; metastasis; human; RT-PCR; ss.
 OS Homo sapiens.
 XX
 XX US5882864-A.
 XX 16-MAR-1999.
 XX 31-JUL-1996; 96US-0692787.
 XX 31-JUL-1995; 95US-0001655.
 XX 31-JUL-1996; 96US-0692787.
 XX (UROC-) UROCOR INC.

| | |
|--------------------------|---|
| CC | prostate breast and bladder tissues and are diagnostic of the potential |
| CC | for metastatic spread of malignant prostate tumours. The nucleic acid can |
| CC | also be used as targets for therapeutic intervention in prostate cancer, |
| CC | benign prostatic hyperplasia (BPH), bladder cancer or breast cancer. The |
| CC | markers may be used to design specific probes and primers, for the rapid |
| CC | analysis of prostate, bladder or breast biopsy samples. The probes and |
| CC | primers may also be used for in situ hybridization or in situ PCR |
| CC | detection and diagnosis. They may also be used to identify and isolate |
| CC | full length gene sequences from various DNA libraries. Antibodies |
| CC | against the polypeptide products of the markers can be used to treat |
| CC | prostate cancer, bladder cancer or breast cancer. The encoded proteins |
| CC | may be used to detect antibodies. The proteins and antibodies can be |
| CC | used in immunodetection methods for detecting or quantifying the cancers, |
| CC | and for clinical diagnosis of these cancers. The antibodies may also be |
| CC | used for radiolabeling to quantify and localize the encoded proteins. |
| SQ | Sequence 673 BP; 222 A; 140 C; 145 G; 166 T; 0 other; |
| XX | |
| XX | |
| Query Match | 1.7%; Score 17; DB 21; Length 673; |
| Best Local Similarity | 100.0%; Pred. No. 65; |
| Matches 17; Conservative | 0; Mismatches 0; Indels 0; Gaps 0; |
| OY | 537 gaaccacacagagaag 553 |
| | |
| Dd | 542 gaaccacacagagaag 558 |
| RESULT 35 | |
| AAZ33534 | |
| ID | AAZ33534 standard; cDNA; 697 BP. |
| XX | |
| AC | AAZ33534; |
| XX | |

| | |
|----|---|
| DE | Human prostate cancer-associated EST 58. |
| XX | |
| KW | Expressed sequence tag; EST; prostate; tumor; treatment; gene therapy; |
| KW | cancer; tissue specificity; human; ss. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| FN | DE19811194-A1. |
| XX | |
| PD | 16-SEP-1999. |
| XX | |
| PF | 10-MAR-1998; 98DE-1011194. |
| XX | |
| PR | 10-MAR-1998; 98DE-1011194. |
| XX | |
| PA | (META-) METAGEN GES GENOMFORSCHUNG MBH. |
| XX | |
| PI | Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A; |
| XX | <u>WPI: 1999-519629/44.</u> |
| DR | P-PSDB; AAY48440. |
| XX | |
| PT | New nucleic acid expressed at high level in normal prostatic tissue and |
| PT | encoded polypeptides, used to treat cancer and screen for therapeutic |
| PT | agents |
| XX | |
| PS | Claim 1a; 119; 194p; German. |
| XX | |
| CC | This invention describes novel nucleic acid sequences (A) that are |
| CC | expressed at high level in normal prostatic tissue. Polypeptides (1) |
| CC | encoded by (A) are used: (a) for identifying agents for treatment of |
| CC | prostatic cancer and (b) for therapy of prostate cancer, optionally |
| CC | where expressed by gene therapy methods. (A) is also used to isolate |
| CC | full-length genes (for gene therapy) and for recombinant production of |
| CC | (1), which can be used to raise specific antibodies. (A) are identified |
| CC | by assembly of ESTs (expressed sequence tags) before these are analyzed |
| CC | for expression pattern (tissue specificity). This approach eliminates |
| CC | |

CC many of the false results, as regards tissue specificity, associated
CC with known methods that use single (usually short) ESTs. AA233477-233540
CC represent expressed sequence tags described in the method of the
CC invention.

XX
SQ Sequence 697 BP; 223 A; 140 C; 196 G; 138 T; 0 other;

Query Match 1.7%; Score 17; DB 20; Length 697;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 aacagcagccagcgccaa 129
IIIIIIIIIIIIIIIIIIII
Db 110 aacagcagccagcgccaa 126

RESULT 36

AAAC39218
ID AAC39218 standard; DNA; 763 BP.

XX AAC39218;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 23810.

XX Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

PD 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0128845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158366.
 PR 13-OCT-1999; 99US-0158293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160778.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160810.
 PR 22-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 1.7%; Score 17; DB 21; Length 763;
 Best Local Similarity 100.0%; Pred No. 64;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 gccgcgtcgtcatccccc 233
 Db 204 gccgcgtcgtcatccccc 220
 <4

RESULT 37
 AAC99839/c
 ID AAC99839 standard; cDNA; 800 BP.
 XX
 AC AAC99839;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DE Human secreted protein gene 22 SEQ ID NO:32.
 XX
 KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
 KW antirheumatic; antiproliferative; cytosolic; cardiac; vasotropic;
 KW cerebroprotective; nootropic; neuroprotective; antibacterial; vitrucide;
 KW fungicide; ophthalmological; gene therapy; pathological condition;
 KW autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;
 KW neoplasm; cardiovascular disorder; cardiac arrest; cerebral ischaemia;
 KW cerebrovascular disorder; angiogenesis; nervous system disorder;
 KW Alzheimer's disease; infection; ocular disorder; corneal infection;
 KW wound healing; skin aging; food additive; preservative;
 KW Chromosome 11; ss.
 KW
 OS Homo sapiens.
 XX
 PN WO200070042-A1.
 XX
 PD 23-NOV-2000.
 XX
 PF 11-MAY-2000. 2000MO-US12788.
 XX
 XX 13-MAY-1999; 99US-0134068.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Moore PA, Young PE, Komatsuolis GA, Birse CE;
 PI Duan RD, Florence KA, Soppet DR;
 XX
 DR WPI: 2000-679828/66.
 DR P-PSDB; AAB56098.
 XX
 PT Isolated nucleic acid molecule encoding a human secreted protein is
 PT used in preventing, treating or ameliorating a medical condition -
 XX
 PS Claim 1; Page 868; 1065pp; English.
 XX
 CC The polynucleotide sequences given in AAC99818 to AAC99977 encode the
 CC human secreted proteins given in AAB56077 to AAB56362. Human secreted
 CC proteins have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: immunosuppressive;
 CC antirheumatic; antiproliferative; antiproliferative; cardiac;
 CC vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;
 CC vitrucide; fungicide; and ophthalmological. The human secreted
 CC polynucleotides and proteins can be used to prevent, treat or ameliorate
 CC a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,
 CC dogs, chickens or sheep. They are also used in diagnosing a pathological
 CC condition or susceptibility to a pathological condition. Disorders which
 CC are diagnosed or treated include autoimmune diseases e.g. rheumatoid
 CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
 CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
 CC disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders
 CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and
 CC fungi and ocular disorders e.g. corneal infection. The proteins can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before

CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The proteins can also be used as a
CC food additive or preservative to increase or decrease storage
CC capabilities. AAC99809 to AAC99817 and AAB56076 represent sequences used
CC in the exemplification of the present invention.

XX Sequence 800 BP; 154 A; 187 C; 154 G; 305 T; 0 other;

Query Match 1.7%; Score 17; DB 21; Length 800;

Best Local Similarity 100.0%; Pred. NO. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 113 aacagcagccagcgccaa 129
|||||
DB 743 AACAGCAGCCAGCCCAA 727

RESULT 38

AAC42507
ID AAC42507 standard; DNA; 1148 BP.

XX AAC42507;

DE 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 35818.

XX Hybridisation assay: genetic mapping; gene expression control;

KM protein identification; signal transduction pathway;

KM metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

PN EPI03405-A2.

XX PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139730.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145226.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147250.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

PR 29-OCT-1999; 99US-0162142.
Query Match 1.7%; Score 17; DB 21; Length 1148;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 222 tgcgtcacccctgcgat 238
|||||
Db 593 tgcgtcacccctgcgat 609
RESULT 39
AAC49984
ID AAC49984 standard; DNA; 1273 BP.
XX
AC AAC49984;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 63167.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
FN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154479.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159684.
PR 18-OCT-1999; 99US-0159588.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.7%; Score 17; DB 21; Length 1273;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 222 tgcgcacccctgcac 238

Db 522 tgcgcacccctgcacat 538
|||||
RESULT 40
AAC37359
ID AAC37359 standard; DNA: 1277 BP.
XX
AC AAC37359;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 17090.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123160.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139464.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145193.
PR 23-JUL-1999; 99US-0145214.
PR 23-JUL-1999; 99US-0145215.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145916.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147304.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.

```

PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-015138.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 1.7%; Score 17; DB 21; Length 1277;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 222 tgctgcatccctcgcat 238
Db 524 tgctgcatccctcgcat 540

```

```

RESULT 41
AAV64526/c
ID AAV64526 standard; DNA; 1338 BP.

```

```

XX XX
AC AAV64526;
XX XX
XX 27-JAN-1999 (first entry)
DT DT
XX M. tuberculosis immunogenic polypeptide XP31 DNA.
XX XX
XX Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
KM vaccine; pharmaceutical; infection; diagnosis; ss.
XX XX
OS Mycobacterium tuberculosis.
XX PN
XX WO9816646-A2.
XX PD
XX 23-APR-1998.
XX PF
XX 07-OCT-1997; 97WO-US18293.
XX PR
XX 13-MAR-1997; 97US-0818112.
XX PR
XX 11-OCT-1996; 96US-0730510.
XX PA
XX (CORI-) CORIXA CORP.
XX PI
XX Campos-Neto A, Dillon DC, Houghton R, Lodes MJ;
XX PI Reed SG, Skelky YAW, Twardzik DR, Vedvick TS;
XX DR
XX WPI; 1998-261042/23.
XX PT
XX Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used
XX PT to develop products for the detection of M. tuberculosis infection
XX PT and for diagnosis, treatment and prevention of tuberculosis
XX PS
XX Claim 31a; Page 162-163; 230pp; English.
XX XX
CC This sequence encodes an immunogenic portion of a soluble Mycobacterium
CC tuberculosis (MT) antigen which can be used in a method for inducing
CC protective immunity against tuberculosis (TB). This sequence can be
CC formulated into vaccines and/or pharmaceutical compositions for
CC immunising against M. tuberculosis infection or may be used for the
CC diagnosis of tuberculosis.
XX XX
SO Sequence 1338 BP; 199 A; 432 C; 485 G; 222 T; 0 other;

```

```

Query Match 1.7%; Score 17; DB 19; Length 1338;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 505 gccagcgagcctcgatc 521
Db 1147 GCCAGCGGCGCTCTGTC 1131

```

```

RESULT 42
AAV44417/c
ID AAV44417 standard; DNA; 1338 BP.
XX AC
XX AAV44417;
XX DT
XX 09-NOV-1998 (first entry)
XX DE
XX Mycobacterium tuberculosis antigen XP31 DNA.
XX XX
XX Tuberculosis; infection; diagnosis; antigen; XP31; ss.
XX OS
XX Mycobacterium tuberculosis strain Erdman.
XX PN
XX WO9816645-A2.
XX PD
XX 23-APR-1998.
XX PF
XX 07-OCT-1997; 97WO-US18214.
XX XX

```

| | | |
|----|---|---------------|
| XX | 13-MAR-1997: | 97US-0818111. |
| PR | 11-OCT-1996; | 96US-0729622. |
| XX | | |
| PA | (CORI-) CORIYA CORP. | |
| XX | | |
| PI | Campos-Meto A, Dillon DC, Houghton R, Lodes MJ; | |
| PI | Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS; | |
| XX | WPI; 1998-251292/72. | |
| DR | | |
| XX | | |
| PT | New isolated Mycobacterium tuberculosis polypeptides and DNA - used | |
| PT | to develop products for the detection of M. tuberculosis infection | |
| XX | and diagnosis of tuberculosis | |
| PT | | |
| PS | Claim 11a; Page 171-172; 250pp; English. | |
| XX | | |
| XX | | |
| CC | This DNA sequence codes for a portion of Mycobacterium tuberculosis | |
| CC | antigen XP31. It was isolated from a M. tuberculosis strain Erdman | |
| CC | genomic DNA expression library using sera from patients having | |
| CC | extrapulmonary tuberculosis. The clone bears some similarity to | |
| CC | known sequences. The invention relates to compositions and methods | |
| CC | for diagnosing tuberculosis. It provides polypeptides (see | |
| CC | AAW64291-W64379) comprising an antigenic portion of a soluble M. | |
| CC | tuberculosis antigen, or an immunogenic portion of an M. | |
| CC | tuberculosis antigen, as well as DNA sequences encoding such | |
| CC | polypeptides, recombinant expression vectors and transformed or | |
| CC | transfected host cells. Also claimed are methods and diagnostic | |
| CC | kits for detecting M. tuberculosis infection in a patient using | |
| CC | these polypeptides, antibodies or oligonucleotide probes and | |
| XX | primers. | |
| XX | | |
| XX | Sequence 1338 BP; 199 A; 432 C; 485 G; 222 T; 0 other; | |

| | | | | |
|--------------------------|--------|-------------------|-------|------------|
| Query Match | 1.7% | Score 17 | DB 19 | Length 138 |
| Best Local Similarity | 100.0% | Pred. No. 64 | | |
| Matches 17, Conservative | 0 | Mismatches | 0 | Indels 0 |
| Gaps | 0 | | | |
| QY | 505 | gccacgcgcgcctcgtc | 521 | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| Db | 1147 | gccacgcgcgcctcgtc | 1131 | |

| | | |
|------------|--|--|
| XX | | |
| PI | Campes-Neto A, Dillon DC, Hendrickson RC, Houghton R; | |
| PI | Lodes MU, Reed SG, Skeiky YAW, Twardzik DR, Vedralick TS; | |
| XX | | |
| PA | (CORI-) CORIXA CORP. | |
| XX | | |
| FR | 05-MAY-1998; 98US-0072667. | |
| PR | 18-FEB-1998; 98US-0025197. | |
| XX | | |
| PF | 17-FEB-1999; 99MO-US03268. | |
| XX | | |
| PD | 26-AUG-1999. | |
| XX | | |
| PN | WO9942076-A2. | |
| OS | Mycobacterium tuberculosis. | |
| XX | | |
| KM | Mycobacterium tuberculosis; M. tuberculosis: antigen; immunogen; | |
| KW | immunotherapy; diagnosis; immunisation; vaccine; infection; | |
| KM | immune response; skin test; ss. | |
| XX | | |
| DE | | |
| XX | | |
| DT | M. tuberculosis antigen XP31 DNA sequence. | |
| AC | | |
| XX | | |
| DT | 05-NOV-1999 (first entry) | |
| AC | | |
| XX | | |
| ID | AAZ19327 standard; DNA; 1338 BP. | |
| XX | | |
| XX | | |
| AC | AAZ19327; | |
| XX | | |
| RESULT | 43 | |
| AAZ19327/C | | |

| | |
|----|---|
| DR | WPI; 1995-527409/44. |
| XX | |
| PT | New antigens from Mycobacterium tuberculosis useful in diagnostic |
| PT | skin tests and protective or therapeutic vaccines or compositions |
| XX | |
| XX | Claim 11; Page 157-158; 299pp; English. |
| PS | |
| XX | |
| CC | The present invention describes polypeptides comprising an immunogenic |
| CC | part of a Mycobacterium tuberculosis antigen (Ag). Also described |
| CC | are vaccines and fusion protein containing M. tuberculosis Ag's. |
| CC | M. tuberculosis Ag's, DNAs encoding them, derived fusion proteins and |
| CC | other polypeptides fragments, can be used in pharmaceutical composition |
| CC | or vaccines to generate a protective or therapeutic immune response to |
| CC | M. tuberculosis and as reagents in skin tests for diagnosis of |
| CC | tuberculosis. Ag can induce proliferation of, or cytokine secretion |
| CC | by, T, B or natural killer cells and/or macrophages in |
| CC | tuberculosis-immune subjects. AA219249 and AA219460 and AA219083 to |
| CC | AA219225 are used in the exemplification of the present invention. |
| XX | |
| XX | |
| XX | Sequence 1338 BP; 199 A; 432 C; 465 G; 222 T; 0 other; |

| | Query Match | 1.7% | Score 17 | DB 20 | Length 138 |
|----|--------------------------|----------------|--------------|----------|------------|
| | Best Local Similarity | 100.0% | Pred. No. 64 | | |
| | Matches 17 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 |
| QY | 505 gccagcgcgcgcctcgtc | 521 | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| Db | 1147 gccacggcggcgcctcgtc | 1131 | | | |

| | | |
|----|--|--|
| CC | RESULT 44 | |
| CC | AAZ19115/C | |
| CC | AAZ19115 standard; DNA; 1338 BP. | |
| CC | AAZ19115; | |
| CC | 05-NOV-1999 (first entry) | |
| CC | M. tuberculosis recombinant antigen DNA encoding XP3L. | |
| CC | Antigen; diagnosis; detection; infection; antibody; immunisation; | |
| CC | vaccine; immunity; ss. | |
| CC | Mycobacterium tuberculosis. | |
| CC | WO9942118-A2. | |
| CC | 26-AUG-1999. | |
| CC | 17-FEB-1999; 99WO-US03265. | |
| CC | 05-MAY-1998; 98US-0072596. | |
| CC | 18-FEB-1998; 98US-0024753. | |
| CC | (CORI-) CORIXA CORP. | |
| CC | Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R; | |
| CC | Lodes MJ, Reed SG, Skeiky YAM, Twardzik DR, Vedvick TS; | |
| CC | WPI; 1999-527416/44. | |
| CC | New polypeptide comprising antigenic portions of M. tuberculosis | |
| CC | Claim 11a; Page 202-203; 323pp; English. | |
| CC | This invention describes novel recombinant antigens and their encoding | |
| CC | nucleic acids derived from Mycobacterium tuberculosis. The novel | |
| CC | polypeptides are useful for detecting M. tuberculosis infection in a | |
| CC | biological sample by detecting antibodies which bind with the | |
| CC | polypeptides, and are useful as vaccines for immunizing against | |
| CC | M. tuberculosis infection. The new detection methods are needed as | |
| CC | current vaccination strategies do not provide 100% immunity. | |

CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
 CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
 CC disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders
 CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and
 CC fungi and ocular disorders e.g. corneal infection. The polypeptides can
 CC also be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides are used to
 CC modulate mammalian metabolism, to change mammal's mental state or
 CC physical state by influencing biorhythms circadian rhythms, depression
 CC tendency for violence tolerance for pain, reproductive capabilities,
 CC hormonal or endocrine levels, appetite, libido, memory, stress or other
 CC cognitive qualities, as a food additive or preservative, such as to
 CC increase or decrease storage capabilities, fat content, lipid, protein,
 CC carbohydrates, vitamins, minerals, cofactors or other nutritional
 CC components.

CC Sequence 1647 BP; 425 A; 301 C; 329 G; 583 T; 9 other;

Query Match

Best Local Similarity 1.7%; Score 17; DB 21; Length 1647;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 485 ttgtgcgcagcctaagt 501
 ||||||||||||||||
 Db 1018 ttgtgcgcagcctaagt 1034

RESULT 47

AAA26361
 ID AAA26361 standard; cDNA; 1947 BP.

XX AAA26361;

XX 29-JUN-2000 (first entry)

XX Human secreted protein gene 16 SEQ ID NO:26.

XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 KW antiHIV; antiinflammatory; neurotropic; neuroprotective; antiallergic;
 KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;
 KW antipruritic; cardiant; gene therapy; cancer; neurological disorder;
 KW immune disease; inflammation; blood disorder; tumour; chromosome 20; ss.

XX Homo sapiens.

XX WO200006698-A1.

XX 10-FEB-2000.

XX 29-JUL-1999; 99WO-US17130.

XX 30-JUL-1998; 98US-0094657.

XX 05-AUG-1998; 98US-0095486.

XX 06-AUG-1998; 98US-0095454.

XX 12-AUG-1998; 98US-0096319.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
 PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
 PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;

XX WPI; 2000-195282/17.

XX P-PSDB; AAY91466.

XX New isolated human genes and the secreted polypeptides they encode,
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders

PS Claim 1; Page 382-383; 634pp; English.

XX The polynucleotide sequences given in AAA26346 to AAA26458 encode the
 CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
 CC proteins can have activities based on the tissues and cells they are
 CC expressed in. Examples of the activities are: cytostatic;
 CC immunosuppressive; antiHIV; antiinflammatory; neurotropic; neuroprotective;
 CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
 CC antiasthma; antipruritic; and cardiant. The polynucleotides and their
 CC corresponding secreted proteins are useful for preventing, treating or
 CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
 CC pathological conditions can be diagnosed by determining the amount of the
 CC proteins in a sample or by determining the presence of mutations in the
 CC polynucleotides. Specific uses are described for each of the
 CC polynucleotides, based on which tissues they are most highly expressed
 CC in, and include developing products for the diagnosis or treatment of
 CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
 CC and foetal deficiencies, blood disorders, diseases of the immune system,
 CC autoimmune diseases, hepatic and renal disease, inflammation,
 CC allergies, Alzheimer's and behavioural disorders, schizophrenia,
 CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
 CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
 CC cardiovascular disorders, reproductive disorders, gastrointestinal
 CC disorders, respiratory disorders and metabolic disorders. The proteins
 CC or polynucleotides can also be used as food additives or preservatives.
 CC The proteins are also useful for identifying their binding partners.
 CC AAA26337 to AAA26345 and AAY91450 are sequences used in the
 CC exemplification of the present invention.

XX Sequence 1947 BP; 534 A; 364 C; 434 G; 615 T; 0 other;

XX Query Match 1.7%; Score 17; DB 21; Length 1947;

XX Best Local Similarity 100.0%; Pred. NO. 63;
 XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 655 ttgtgcgcagcagcata 671
 ||||||||||||||||
 Db 1828 ttgtgcgcagcagcata 1844

RESULT 48

AAA91185
 ID AAA91185 standard; DNA; 2459 BP.

XX AAA91185;

XX 08-MAY-2001 (first entry)

XX Zea mays ZMEIN3-2 coding sequence.

XX Zm EIN3-1; Zm EIN3-2; corn; ethylene signal transduction pathway; maize;
 KW ethylene-mediated response; crowding tolerance; seed set; maturation;
 KW seed development; growth in impacted soil; flooding tolerance;
 KW senescence; disease resistance; ds.

XX Zea mays.

XX Key Location/Qualifiers

XX CDS 285..2228

XX FT /tag= a

XX PN /product= "Zm EIN3-3"

XX WO200105953-A2.

XX 25-JAN-2001.

XX 07-JUL-2000; 2000WO-US18746.

XX 15-JUL-1999; 99US-0143965.
 XX (PION-) PIONEER HI-BRED INT INC.

PI Stimmans CR:
 XX WPI: 2001-147338/15.
 DR P-PSDB; AAY97667.
 XX
 PT Novel maize ethylene signalling pathway EIN3 gene useful for modulating
 PT the level of EIN3 in maize plants, including crowding tolerance, growth
 PT in impacted soils, flooding tolerance and disease resistance
 PS
 PS Claim 1: Page 82-85; 86pp; English.
 XX
 CC This sequence encodes the corn zm EIN3-2 protein. The protein is involved
 CC in the ethylene signal transduction pathway, and is an EIN3 homologue.
 CC An expression cassette comprising the DNA sequence is useful for
 CC modulating the level of EIN3 in a plant, in particular a maize plant. The
 CC maize genes are nuclear transcription factors that promote the
 CC ethylene-mediated responses, including crowding tolerance, seed set and
 CC development, growth in impacted soils, flooding tolerance, maturation and
 CC senescence and disease resistance. Diminishment of ethylene action in
 CC plant, in particular cereals such as maize, by reducing the expression or
 CC activity of the DNA promotes tolerance to close spacing with reduced
 CC stress and yield loss. The DNA is useful as a probe or amplification
 CC primer in the detection, quantitation or isolation of gene transcripts,
 CC in detecting deficiencies in the level of mRNA in screening for desired
 CC transgenic plants, for detecting mutations in gene, for monitoring
 CC upregulation of expression or changes in enzyme activity in screening
 CC assays, orthologs, or paralog of the gene, or for site directed
 CC mutagenesis in eukaryotic cells. The nucleic acid can also be used for
 CC recombinant expression of polypeptides or as immunogens in the
 CC preparation and/or screening of antibodies. The proteins can be employed
 CC in assays for enzyme agonists or antagonists of enzyme function, or as
 CC immunogens or antigens to obtain antibodies specifically immunoreactive
 CC with the protein. Plants expressing the DNA germinate better in compacted
 CC soils and in flooded conditions or water-logged soils, resulting in
 CC higher stand counts.
 XX
 CC Sequence 2459 BP; 587 A; 747 C; 696 G; 429 T; 0 other:
 SQ
 Query Match 1.7%; Score 17; DB 22; Length 2459;
 Best Local Similarity 100.0%; Pred. No. 63;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 817 gagctgcagcagagccg 833
 ||||||||||||||||
 Db 570 gagctgcagcagagccg 586
 RESULT 49
 AAX14998/C
 ID AAX14998 standard; DNA; 2545 BP.
 XX
 AC AAX14998;
 XX
 DT 15-APR-1999 (first entry)
 XX
 DE DNA encoding a monokine induced by gamma-interferon (MIG).
 XX
 KW Monokine induced by gamma-interferon; MIG; CXK chemokine; metastasis;
 KW angiogenesis inhibitor; angiotensin inducer; tumour growth inhibition;
 KW haemangiomas; rheumatoid arthritis; atherosclerosis; meningioma;
 KW idiopathic pulmonary fibrosis; benign prostatic hypertrophy; psoriasis;
 KW vascular restenosis; arteriovenous malformation; neovascular glaucoma;
 KW angiofibroma; haemophilic joint; hypertrophic scar; Osler-Weber syndrome;
 KW pyogenic granuloma retrolental fibroplasia; scleroderma; trachoma;
 KW vascular adhesion; synovitis; dermatitis; endometriosis; pterygium;
 KW diabetic retinopathy; neovascularisation; chronic bronchitis;
 KW adult respiratory distress syndrome; ARDS; pseudogout;
 KW cystic fibrosis; ss.
 XX
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers

FT CDS 40..417
 FT /*tag= a
 FT /product= MIG
 XX
 PN US5871723-A.
 XX
 PD 16-FEB-1999.
 XX
 PF 06-JUN-1995; 95US-0468819.
 XX
 PR 06-JUN-1995; 95US-0468819.
 XX
 PA (UNMI) UNIV MICHIGAN.
 PI Kunkel SL, Polymer PJ, Strieter RW;
 DR WPI: 1999-165569/14.
 DR P-PSDB; AAW96710.
 XX
 PT Use of chemokines with a conserved Cys Xaa Cys (CXC) sequence -
 PT which do not contain amino acid sequence ELR, for inhibiting
 PT angiogenesis in tumours, rheumatoid arthritis, restenosis or
 PT glaucoma
 XX
 PS Example 13: Columns 123-126; 145pp; English.
 XX
 CC The present sequence encodes a monokine induced by gamma-interferon
 CC (MIG). MIG is a CXK chemokine that is an inhibitor of angiogenesis.
 CC The specification describes methods for inhibiting angiogenesis or for
 CC inducing angiostasis, using chemokines (with a conserved Cys Xaa Cys
 CC (CXC) sequence at the N-terminal) other than platelet factor-4, and
 CC which do not contain the amino acid sequence ELR. The methods are useful
 CC for inhibiting tumour growth and metastasis and for treating diseases
 CC such as haemangiomas, rheumatoid arthritis, atherosclerosis and
 CC idiopathic pulmonary fibrosis (IPF), benign prostatic hypertrophy (BPH),
 CC vascular restenosis, arteriovenous malformations (AVM), meningioma,
 CC neovascular glaucoma, psoriasis, angiofibroma, haemophilic joints,
 CC hypertrophic scars, Osler-Weber syndrome, pyogenic granuloma retrolental
 CC fibroplasia, scleroderma, trachoma, vascular adhesions, synovitis,
 CC dermatitis, endometriosis, pterygium, diabetic retinopathy
 CC neovascularisation associated with corneal injury or grafts, adult
 CC respiratory distress syndrome (ARDS), chronic bronchitis, pseudogout
 CC and cystic fibrosis.
 XX
 CC Sequence 2545 BP; 755 A; 581 C; 457 G; 752 T; 0 other:
 SQ
 Query Match 1.7%; Score 17; DB 20; Length 2545;
 Best Local Similarity 100.0%; Pred. No. 63;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 787 ctgaactgggttaga 803
 ||||||||||||||||
 Db 277 CTGATCTGGGTTTACA 261
 RESULT 50
 AAA74878/C
 ID AAA74878 standard; DNA; 2545 BP.
 XX
 AC AAA74878;
 XX
 DT 17-JAN-2001 (first entry)
 XX
 DE Human chemokine coding sequence spg ID NO: 39.
 XX
 KW Macrophage recruitment; chemokine derivative; MCP-1; osteoporosis;
 KW monocyte chemoattractant protein-1; inflammation; atherosclerosis; HIV;
 KW AIDS; stroke; psoriasis; autoimmune disease; hypertension; endotoxaemia;
 KW basophil-mediated disease; myocardial infarction; acute ischaemia;
 KW rheumatoid arthritis; contraception; ds.
 XX
 XX Homo sapiens.
 OS

| Key | Location/Qualifiers |
|---|--|
| FT CDS | 40..417 |
| FT | /tag= a |
| FT | /product= "human chemokine" |
| PN | MO200042071-A2. |
| PD | 20-JUL-2000. |
| PF | 12-JAN-2000; 2000WO-US00821. |
| PR | 12-JAN-1999; 99US-0229071. |
| PR | 17-MAR-1999; 99US-0271192. |
| PR | 01-DEC-1999; 99US-0452406. |
| PA | (NEOR-) NEORX CORP. |
| PI | Gralinger DJ, Tatalick LM; |
| DR | WPI: 2000-499101/44. |
| DR | P-PSDB; AAB15803. |
| XX | New peptide 3, amide and heterocyclic compounds and saccharide |
| PT | conjugates used for inhibiting chemokine induced activity and for |
| PT | treating e.g. stroke, vascular diseases, autoimmune diseases and tumour |
| PT | growth - |
| XX | Disclosure: Page 356-358; 387pp; English. |
| PS | |
| XX | |
| CC | The present invention concerns the identification of a number of |
| CC | chemokines which can be used to produce derivatives, agonists and |
| CC | antagonists which are then useful in disease treatment. The chemokines |
| CC | include sequences AAB15785-B15794, AAB15803-B15813 and AAB15831-B15848. |
| CC | These chemokine derivatives can be used to treat diseases such as |
| CC | autoimmune diseases, atherosclerosis, osteoporosis, HIV infection and |
| CC | AIDS, psoriasis, inflammatory diseases, hypertension, basophil-mediated |
| CC | diseases, endotoxaemia, myocardial infarction, acute ischaemia and |
| CC | rheumatoid arthritis, and can be used to prevent strokes and as |
| CC | contraceptives. The chemokine coding sequences AA47488-A7488 can be |
| CC | used in gene therapy for the same diseases, as well as in the production |
| CC | of animal models. |
| XX | |
| CC | |
| XX | |
| SO | Sequence 2545 BP; 755 A; 581 C; 457 G; 752 T; 0 other; |
| Query Match | 1.7%; Score 17; DB 21; Length 2545; |
| Best Local Similarity | 100.0%; Pred. No. 63; |
| Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| 07 787 ctgaatctggtttaga 803 | |
| | |
| 277 ctgaatctggtttaga 261 | |

Search completed: September 21, 2001, 16:27:43
Job time: 9633 sec

THIS PAGE BLANK (USPTO)

117: qb_est148:*
118: qb_est149:*
119: qb_est150:*
120: qb_est151:*
121: qb_est152:*
122: qb_est153:*
123: qb_est154:*
124: qb_est155:*
125: qb_est156:*
126: qb_est157:*
127: qb_est158:*
128: qb_est159:*
129: qb_est160:*
130: qb_est161:*
131: qb_est162:*
132: qb_est163:*
133: qb_est164:*
134: qb_est165:*
135: qb_est166:*
136: qb_est167:*
137: qb_est168:*
138: qb_est169:*
139: qb_est170:*
140: qb_est171:*
141: qb_est172:*
142: qb_est173:*
143: qb_est174:*
144: qb_est175:*
145: qb_est176:*
146: qb_est177:*
147: qb_est178:*
148: qb_est179:*
149: qb_est180:*
150: qb_est181:*
151: qb_est182:*
152: qb_est183:*
153: qb_est184:*
154: qb_est185:*
155: qb_est186:*
156: qb_est187:*
157: qb_est188:*
158: qb_est189:*
159: qb_est190:*
160: qb_est191:*
161: qb_est192:*
162: qb_est193:*
163: qb_est194:*
164: qb_est195:*
165: qb_est196:*
166: qb_est197:*
167: qb_est198:*
168: qb_est199:*
169: qb_est200:*
170: qb_est201:*
171: qb_est202:*
172: qb_est203:*
173: qb_est204:*
174: qb_est205:*
175: qb_est206:*
176: qb_est207:*
177: qb_est208:*
178: qb_est209:*
179: qb_est210:*
180: qb_est211:*
181: qb_est212:*
182: qb_est213:*
183: qb_est214:*
184: qb_est215:*
185: qb_est216:*
186: qb_est217:*
187: qb_est218:*
188: qb_est219:*
189: qb_est220:*

190: qb_est110:*
191: qb_est111:*
192: qb_hic:*
193: em_gss_fun1:*
194: em_gss_fun2:*
195: em_gss_fun3:*
196: em_gss_fun4:*
197: em_gss_fun5:*
198: em_gss_fun6:*
199: em_gss_fun7:*
200: em_gss_fun8:*
201: em_gss_fun9:*
202: em_gss_fun10:*
203: em_gss_fun11:*
204: em_gss_fun12:*
205: em_gss_fun13:*
206: em_gss_fun14:*
207: em_gss_fun15:*
208: em_gss_fun16:*
209: em_gss_fun17:*
210: em_gss_fun18:*
211: em_gss_fun19:*
212: em_gss_fun20:*
213: em_gss_fun21:*
214: em_gss_fun22:*
215: em_gss_fun23:*
216: em_gss_fun24:*
217: em_gss_fun25:*
218: em_gss_fun26:*
219: em_gss_fun27:*
220: em_gss_fun28:*
221: em_gss_fun29:*
222: em_gss_fun30:*
223: em_gss_fun31:*
224: em_gss_fun32:*
225: em_gss_fun33:*
226: em_gss_fun34:*
227: em_gss_fun35:*
228: em_gss_fun36:*
229: em_gss_fun37:*
230: em_gss_fun38:*
231: em_gss_fun39:*
232: em_gss_fun40:*
233: em_gss_fun41:*
234: em_gss_fun42:*
235: em_gss_fun43:*
236: em_gss_fun44:*
237: em_gss_fun45:*
238: em_gss_fun46:*
239: em_gss_fun47:*
240: em_gss_fun48:*
241: em_gss_fun49:*
242: em_gss_fun50:*
243: em_gss_fun51:*
244: em_gss_fun52:*
245: em_gss_fun53:*
246: em_gss_fun54:*
247: em_gss_fun55:*
248: em_gss_fun56:*
249: em_gss_fun57:*
250: em_gss_fun58:*
251: em_gss_fun59:*
252: em_gss_fun60:*
253: em_gss_fun61:*
254: em_gss_fun62:*
255: em_gss_fun63:*
256: em_gss_fun64:*
257: em_gss_fun65:*
258: em_gss_fun66:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|----------------------|
| 1 | 2.2 | 894 | 154 | BG481731 | BG481731 602528280 |
| 2 | 2.0 | 222 | 10 | AA645779 | AA645779 vnl0h04..r |
| 3 | 2.0 | 414 | 113 | AM268708 | AM268708 xv35b05.x |
| 4 | 2.0 | 428 | 24 | A1769658 | A1769658 wj25d02.x |
| 5 | 2.0 | 456 | 15 | A1028162 | A1028162 ov90d02..x |
| 6 | 2.0 | 483 | 113 | AM269670 | AM269670 xv36d11.x |
| 7 | 2.0 | 480 | 22 | AI594177 | AI594177 vnl0h04..y |
| 8 | 2.0 | 917 | 141 | BBE93506 | BBE93506 601438060 |
| 9 | 1.9 | 468 | 229 | AO511509 | AO511509 HS_5074.B |
| 10 | 1.9 | 523 | 227 | AO337527 | AO337527 HS_5017.A |
| 11 | 1.9 | 656 | 245 | A2497878 | A2497878 IM0334624 |
| 12 | 1.9 | 657 | 235 | AO938059 | AO938059 NB6-908R |
| 13 | 1.9 | 713 | 221 | CNS04812 | AI325737 Tetraodon |
| 14 | 1.9 | 747 | 18 | AI325737 | AI325737 mm95d03..y |
| 15 | 1.9 | 842 | 247 | A2668255 | A2668255 ENTKX52TF |
| 16 | 1.9 | 891 | 221 | CNS0370E | AL266639 Tetraodon |
| 17 | 1.8 | 235 | 124 | BB043716 | BB043716 BB043716 |
| 18 | 1.8 | 241 | 28 | AV292150 | AV292150 AV292150 |
| 19 | 1.8 | 244 | 31 | AV526879 | AV526879 AV526879 |
| 20 | 1.8 | 261 | 6 | AA358477 | AA358477 EST67411 |
| 21 | 1.8 | 264 | 31 | AV527317 | AV527317 AV527317 |
| 22 | 1.8 | 294 | 191 | Z29054 | Z29054 ATTS2067 OF |
| 23 | 1.8 | 301 | 26 | AV208201 | AV208201 AV208201 |
| 24 | 1.8 | 312 | 2 | AA095509 | AA095509 14744..seq |
| 25 | 1.8 | 318 | 191 | Z18479 | Z18479 ATTS0855 AC |
| 26 | 1.8 | 327 | 122 | AM896020 | AM896020 QV4-NN003 |
| 27 | 1.8 | 331 | 188 | T121227 | T121227 3235.Lambda |
| 28 | 1.8 | 334 | 104 | AI991901 | AI991901 ws42b06..x |
| 29 | 1.8 | 334 | 119 | AM682518 | AM682518 EST01367 |
| 30 | 1.8 | 340 | 150 | BE521901 | BE521901 UI-R-BTO- |
| 31 | 1.8 | 342 | 136 | BE520970 | BE520970 M16A2STM |
| 32 | 1.8 | 365 | 31 | AV527474 | AV527474 AV527474 |
| 33 | 1.8 | 368 | 189 | T42951 | T42951 6214.Lambda |
| 34 | 1.8 | 370 | 191 | Z47675 | Z47675 ATTS4531 St |
| 35 | 1.8 | 376 | 30 | AV430174 | AV430174 AV430174 |
| 36 | 1.8 | 391 | 31 | AV526330 | AV526330 AV526330 |
| 37 | 1.8 | 392 | 157 | N65689 | N65689 20729.Lambda |
| 38 | 1.8 | 392 | 157 | D78804 | D78804 HNM516H08B |
| 39 | 1.8 | 392 | 226 | AO253544 | AO253544 HS_2227.A |
| 40 | 1.8 | 398 | 31 | AV527255 | AV527255 AV527255 |
| 41 | 1.8 | 404 | 31 | AV526564 | AV526564 AV526564 |
| 42 | 1.8 | 441 | 188 | T21720 | T21720 3728.Lambda |
| 43 | 1.8 | 443 | 158 | H77242 | H77242 17673.Lambda |
| 44 | 1.8 | 446 | 225 | AO228766 | AO228766 HS_2020.A |
| 45 | 1.8 | 461 | 159 | N65546 | N65546 20586.Lambda |
| 46 | 1.8 | 463 | 104 | AI991951 | AI991951 ws43d04.x |
| 47 | 1.8 | 465 | 151 | BE612300 | BE612300 daai7a11.. |
| 48 | 1.8 | 465 | 140 | BE611867 | BE611867 PMO-AN003 |
| 49 | 1.8 | 474 | 244 | A2462258 | A2462258 IM0269023 |
| 50 | 1.8 | 476 | 20 | AI466074 | AI466074 vvs39d02..y |
| 51 | 1.8 | 482 | 234 | AO847780 | AO847780 LMAJFV1..1 |
| 52 | 1.8 | 487 | 234 | AO849674 | AO849674 LMAJFV1..1 |
| 53 | 1.8 | 489 | 18 | AA395581 | AA395581 27378.Lam |
| 54 | 1.8 | 489 | 166 | BE351801 | BE351801 894051G09 |
| 55 | 1.8 | 493 | 164 | BE215806 | BE215806 HV_CEB000 |
| 56 | 1.8 | 495 | 31 | AV527484 | AV527484 AV527484 |
| 57 | 1.8 | 498 | 136 | BE487307 | BE487307 175980.BA |
| 58 | 1.8 | 499 | 166 | BE351915 | BE351915 894053D01 |
| 59 | 1.8 | 500 | 31 | AV527890 | AV527890 AV527890 |
| 60 | 1.8 | 502 | 31 | AV549916 | AV549916 AV549916 |
| 61 | 1.8 | 504 | 14 | AA582948 | AA582948 ua22h06..r |
| 62 | 1.8 | 504 | 14 | AA582948 | AA582948 ua22h06..r |
| 63 | 1.8 | 505 | 188 | T22194 | T22194 4202.Lambda |
| 64 | 1.8 | 512 | 164 | BE215810 | BE215810 HV_CEB000 |
| 65 | 1.8 | 512 | 228 | AO399589 | AO399589 mgx0d018E |
| 66 | 1.8 | 516 | 188 | R64762 | R64762 1336.Lambda |
| 67 | 1.8 | 525 | 188 | T21563 | T21563 3571.Lambda |
| 68 | 1.8 | 527 | 112 | AM168137 | AM168137 xg60d04..x |
| 69 | 1.8 | 529 | 159 | N38270 | N38270 19497.Lambda |
| 70 | 1.8 | 531 | 31 | AV528061 | AV528061 AV528061 |
| 71 | 1.8 | 536 | 16 | AI158964 | AI158964 u142g05..x |
| 72 | 1.8 | 539 | 104 | AI995890 | AI995890 701549071 |
| 73 | 1.8 | 556 | 174 | BG147743 | BG147743 mab53h05. |
| 74 | 1.8 | 565 | 31 | AV567359 | AV567359 AV567359 |
| 75 | 1.8 | 570 | 108 | AU164787 | AU164787 AU164787 |
| 76 | 1.8 | 583 | 31 | H36018 | H36018 14540.Lambda |
| 77 | 1.8 | 592 | 158 | H36018 | H36018 14540.Lambda |
| 78 | 1.8 | 592 | 158 | H76069 | H76069 17774.Lambda |
| 79 | 1.8 | 598 | 151 | BF620011 | BF620011 HVSMEC000 |
| 80 | 1.8 | 599 | 1 | AA042763 | AA042763 25019.CD4 |
| 81 | 1.8 | 618 | 107 | AU101340 | AU101340 AU101340 |
| 82 | 1.8 | 621 | 242 | AZ351625 | AZ351625 1M0089G14 |
| 83 | 1.8 | 682 | 165 | BE291467 | BE291467 601085211 |
| 84 | 1.8 | 700 | 162 | BE037926 | BE037926 AA06E04.A |
| 85 | 1.8 | 707 | 152 | BG323623 | BG323623 602422094 |
| 86 | 1.8 | 711 | 247 | AZ631126 | AZ631126 1M0485G09 |
| 87 | 1.8 | 712 | 251 | AZ687529 | AZ687529 2M0178H23 |
| 88 | 1.8 | 722 | 141 | BE892353 | BE892353 601433812 |
| 89 | 1.8 | 722 | 162 | BE038966 | BE038966 AB07H04.A |
| 90 | 1.8 | 727 | 162 | BE038427 | BE038427 AA17D03.A |
| 91 | 1.8 | 731 | 144 | BF120897 | BF120897 601757705 |
| 92 | 1.8 | 754 | 138 | BE689492 | BE689492 uw55c05..y |
| 93 | 1.8 | 807 | 21 | AI528777 | AI528777 mr72a10..x |
| 94 | 1.8 | 862 | 146 | BF267443 | BF267443 HV_CBA001 |
| 95 | 1.8 | 866 | 150 | BF526531 | BF526531 602070839 |
| 96 | 1.8 | 898 | 235 | AO901727 | AO901727 HS_2019..A |
| 97 | 1.8 | 909 | 220 | CNS01XED | AL11604 Tetraodon |
| 98 | 1.8 | 919 | 221 | CNS030E7 | AL255868 Tetraodon |
| 99 | 1.8 | 920 | 154 | BG479157 | BG479157 602526252 |
| 100 | 1.8 | 952 | 222 | CNS057PS | AL340489 Tetraodon |
| 101 | 1.8 | 968 | 220 | BE741337 | BE741337 601594184 |
| 102 | 1.8 | 986 | 220 | CNS03SRQ | AL229103 Tetraodon |
| 103 | 1.8 | 1009 | 221 | CNS0305J | AL255520 Tetraodon |
| 104 | 1.8 | 1056 | 222 | CNS05E8A | AL333451 Tetraodon |
| 105 | 1.8 | 1078 | 106 | AL543701 | AL543701 AL543701 |
| 106 | 1.8 | 1081 | 222 | CNS054X1 | AL321310 Tetraodon |
| 107 | 1.8 | 1101 | 192 | AK006925 | AK006925 Mus muscu |
| 108 | 1.8 | 1144 | 175 | BF118779 | BF118779 601755194 |
| 109 | 1.8 | 1481 | 145 | BG250472 | BG250472 602362637 |
| 110 | 1.7 | 72 | 158 | H88159 | H88159 yv18c05..s1 |
| 111 | 1.7 | 122 | 159 | N70934 | N70934 za33f06..s1 |
| 112 | 1.7 | 136 | 158 | H59904 | H59904 yrc04e09..s1 |
| 113 | 1.7 | 164 | 23 | AI701300 | AI701300 wcs3907..x |
| 114 | 1.7 | 165 | 1 | AA002343 | AA002343 mg42a03..r |
| 115 | 1.7 | 206 | 1 | AA042040 | AA042040 m306a07..r |
| 116 | 1.7 | 207 | 28 | AV278443 | AV278443 AV278443 |
| 117 | 1.7 | 208 | 174 | BG158928 | BG158928 RH122..42 |
| 118 | 1.7 | 212 | 122 | AM896035 | AM896035 QV4-NN003 |
| 119 | 1.7 | 219 | 238 | AZ083101 | AZ083101 RPEC1-23-4 |
| 120 | 1.7 | 226 | 8 | AA473731 | AA473731 vgs9f03..r |
| 121 | 1.7 | 230 | 191 | Z41347 | Z41347 HSCGR112 n |
| 122 | 1.7 | 240 | 26 | AV196116 | AV196116 AV196116 |
| 123 | 1.7 | 241 | 124 | AM647187 | AM647187 EST325730 |
| 124 | 1.7 | 241 | 155 | C05473 | C05473 C05473 Humna |
| 125 | 1.7 | 242 | 121 | AM809963 | AM809963 MK4-ST012 |
| 126 | 1.7 | 244 | 144 | BF116745 | BF116745 u91f04..y |
| 127 | 1.7 | 244 | 148 | BF730909 | BF730909 mab78f07.. |
| 128 | 1.7 | 248 | 189 | T84122 | T84122 yd33e12..r1 |
| 129 | 1.7 | 248 | 7 | AA416280 | AA416280 yd18b10..s |
| 130 | 1.7 | 251 | 187 | N87803 | N87803 K0130F Humna |
| 131 | 1.7 | 251 | 13 | AA887431 | AA887431 c350c04..s |
| 132 | 1.7 | 253 | 13 | AA741155 | AA741155 n201h02..s |
| 133 | 1.7 | 255 | 11 | AV260749 | AV260749 AV260749 |
| 134 | 1.7 | 260 | 141 | BBE60963 | BBE60963 UI-M-ANI- |
| 135 | 1.7 | 264 | 6 | AA382428 | AA382428 EST95633 |
| 136 | 1.7 | 265 | 161 | BB563790 | BB563790 BB563790 |
| 137 | 1.7 | 266 | 6 | AA383032 | AA383032 EST96333 |
| 138 | 1.7 | 266 | 162 | BB604316 | BB604316 BB604316 |
| 139 | 1.7 | 267 | 11 | AA741789 | AA741789 Lmtv39p3/ |
| 140 | 1.7 | 267 | 11 | AA741789 | AA741789 Lmtv39p3/ |

579 17 1.7 979 173 BG106112
580 17 1.7 980 141 BE903833
581 17 1.7 980 172 BG030132
582 17 1.7 983 220 CNS0347U
583 17 1.7 990 220 BF215841
584 17 1.7 996 220 CNS03861
585 17 1.7 998 220 CNS0360B
586 17 1.7 1001 106 AL561843
587 17 1.7 1015 192 AK002864
588 17 1.7 1021 192 AK003834
589 17 1.7 1022 190 W14751
590 17 1.7 1025 152 BG329221
591 17 1.7 1028 221 CNS03LAK
592 17 1.7 1032 175 BG249520
593 17 1.7 1039 220 CNS02GAL
594 17 1.7 1083 222 CNS05H7M
595 17 1.7 1084 222 CNS04XJS
596 17 1.7 1087 175 BG259333
597 17 1.7 1101 219 CNS000ND
598 17 1.7 1101 219 CNS00G3V
599 17 1.7 1101 222 CNS05A0Z
600 17 1.7 1126 150 BF584626
601 17 1.7 1127 190 W29158
602 17 1.7 1144 141 BE891148
603 17 1.7 1146 174 BG113345
604 17 1.7 1195 145 BF208266
605 17 1.7 1196 192 AK006059
606 17 1.7 1319 167 BE420590
607 17 1.7 1418 153 BG393682
608 17 1.7 1593 192 AK017315
609 17 1.7 1607 146 BF237917
610 17 1.7 1690 150 BF572286
611 17 1.7 1889 173 BG109738
612 16 1.6 559 219 AG022982
613 16 1.6 107 249 AZ761435
614 16 1.6 113 162 T81602
615 16 1.6 121 162 BE028205
616 16 1.6 138 251 AZ922798
617 16 1.6 138 251 AZ922798
618 16 1.6 139 165 BE301846
619 16 1.6 141 133 BA422645
620 16 1.6 142 150 BF534229
621 16 1.6 149 145 BFL15383
622 16 1.6 149 151 BF656411
623 16 1.6 154 251 AZ918325
624 16 1.6 157 188 R84043
625 16 1.6 160 19 A1410706
626 16 1.6 160 161 BB581731
627 16 1.6 165 242 AZ371012
628 16 1.6 168 7 AA422595
629 16 1.6 171 158 H48834
630 16 1.6 171 161 BB600875
631 16 1.6 190 22 A1611757
632 16 1.6 191 164 BE174271
633 16 1.6 192 134 BA448586
634 16 1.6 192 161 BB582796
635 16 1.6 197 162 BE005753
636 16 1.6 202 175 BG243388
637 16 1.6 203 145 BFL19810
638 16 1.6 204 142 BE932425
639 16 1.6 206 161 BB584623
640 16 1.6 207 106 AT002383
641 16 1.6 208 114 AM299903
642 16 1.6 211 27 AV239876
643 16 1.6 211 143 BF006796
644 16 1.6 211 247 AZ634242
645 16 1.6 212 243 AZ411926
646 16 1.6 213 15 A1032589
647 16 1.6 213 25 AV135546
648 16 1.6 215 111 AM125297
649 16 1.6 215 122 AM915765
650 16 1.6 217 108 AV022822
651 16 1.6 217 162 BE011597

BG106112 602289934
BE903833 601677192
BG030132 602297315
AL227091 Tetradon
BF215841 601861357
AL232227 Tetradon
AL230276 Tetradon
AL561843 AL561843
AK002864 Mus muscu
AK003834 Mus muscu
W14751 mb23f09.r1
BG329221 602428263
AL249221 Tetradon
BG259333 602378407
AL066089 Drosophill
AL071401 Drosophill
AL337223 Tetradon
AL311761 Tetradon
BG259333 602378407
BF584626 602098650
W29158 mb9eail.r1
BE891148 601432264
BG113345 602284950
BF208266 601871378
AK006059 Mus muscu
BE420590 HM000.E0
BG393682 602412105
AK017315 Mus muscu
BF237917 601842495
BF572286 602077730
BG109738 602280875
AG022982 Oryz Sat
AZ761435 1M0555F15
T81602 yd30a09.r1
BE028205 ELESTrea57
BF231005 253045.BA
AZ922798 SLC055G10
BE301846 bholc05.y
BA422645 BB422645
BF534229 602047720
BFL15383 QV4-HT053
BF656411 PVL-45.E0
AZ918325 1006004A0
R84043 16002.Lambd
A1410706 EST38999
BB581731 BB581731
AZ371012 1M0122F19
AA422595 vd29e06.s
H48834 yrc34e03.r1
BB600875 BB600875
A1611757 lu78h02.x
BE174271 QV1-HT057
BA448586 BB448586
BB582796 BB582796
BE005753 RC2-BN012
BG243388 602356835
BFL19810 601785791
BE932425 QV3-HT063
BB584623 BB584623
AT002383 AT002383
AM299903 xs44e03.x
AV239876 AV239876
BF006796 1160585.A
AZ634242 1M0489K22
AZ411926 1M0185K12
A1032589 ow70h12.s
AV135546 AV135546
AM125297 UT-M-BH2.
AM915765 EST347069
AV022822 AV022822
BE011597 CM4-BN022

C 652 16 1.6 218 116 AM451421
C 653 16 1.6 220 22 A1623890
C 654 16 1.6 223 171 BF901675
C 655 16 1.6 225 122 AA786141
C 656 16 1.6 225 111 AM087469
C 657 16 1.6 225 112 AM197231
C 658 16 1.6 225 241 AZ290667
C 659 16 1.6 226 103 A1916712
C 660 16 1.6 226 161 BB582536
C 661 16 1.6 227 23 A1652100
C 662 16 1.6 230 29 AV350839
C 663 16 1.6 231 244 AZ479121
C 664 16 1.6 234 251 AZ900753
C 665 16 1.6 235 30 AV418718
C 666 16 1.6 236 222 FRO018439
C 667 16 1.6 240 23 A1683779
C 668 16 1.6 240 249 AZ752706
C 669 16 1.6 241 8 AA477167
C 670 16 1.6 241 147 BF373590
C 671 16 1.6 242 170 BF822858
C 672 16 1.6 242 170 BF823175
C 673 16 1.6 245 244 AZ487110
C 674 16 1.6 246 104 A1978307
C 675 16 1.6 246 161 BB583147
C 676 16 1.6 247 12 AA807648
C 677 16 1.6 247 136 BE516407
C 678 16 1.6 248 101 Z37222
C 679 16 1.6 249 10 AA634103
C 680 16 1.6 252 25 AV154082
C 681 16 1.6 253 248 AZ728927
C 682 16 1.6 256 113 AM214673
C 683 16 1.6 258 228 AO388893
C 684 16 1.6 260 144 BE810749
C 685 16 1.6 260 223 AQ027837
C 686 16 1.6 261 126 BB118069
C 687 16 1.6 261 134 BE528278
C 688 16 1.6 261 156 BG487311
C 689 16 1.6 265 161 BB601852
C 690 16 1.6 267 237 AZ020088
C 691 16 1.6 268 187 R16206
C 692 16 1.6 269 20 A1441353
C 693 16 1.6 272 12 AA825331
C 694 16 1.6 274 130 BB309167
C 695 16 1.6 275 133 BB412379
C 696 16 1.6 276 2 AAO96629
C 697 16 1.6 276 106 AT002342
C 698 16 1.6 276 246 AZ587088
C 699 16 1.6 277 127 BB176196
C 700 16 1.6 279 107 AU029833
C 701 16 1.6 279 121 AM853559
C 702 16 1.6 279 245 AZ500749
C 703 16 1.6 280 109 AV076081
C 704 16 1.6 280 131 AV077408
C 705 16 1.6 280 131 BB319896
C 706 16 1.6 281 161 BB572232
C 707 16 1.6 282 114 AM322270
C 708 16 1.6 282 162 BB603068
C 709 16 1.6 283 190 W63249
C 710 16 1.6 284 172 BG014672
C 711 16 1.6 285 5 AA319807
C 712 16 1.6 285 114 AM293702
C 713 16 1.6 285 172 BG014680
C 714 16 1.6 287 246 AZ625139
C 715 16 1.6 288 128 BB228806
C 716 16 1.6 288 191 Z21331
C 717 16 1.6 288 243 AZ419333
C 718 16 1.6 289 126 BB117978
C 719 16 1.6 289 235 AO922536
C 720 16 1.6 291 132 BB366546
C 721 16 1.6 291 246 AZ588220
C 722 16 1.6 292 113 AM230121
C 723 16 1.6 292 129 BB262861
C 724 16 1.6 292 134 BA452180

AM451421 UT-H-B13-
A1623890 ts24f08.x
BF901675 PM4-MT000
AA786141 j7b07a1.f
AM087469 xb99g10.x
AM197231 xm38d07.x
AZ290667 RPCI-23-1
A1916712 ts80e12.x
BB582536 BB582536
A1652100 wb27h10.x
AV350839 AV350839
AZ479121 1M0299M04
AZ900753 RPCI-24-1
AV418718 AV418718
A1011335 F.rubripie
A1683779 LV53h11.x
AA477167 zud1c08.r
AA477167 zud1c08.r
BF373590 MR0-PT017
BF822858 MR1-RT007
BF823175 MR1-RT007
AZ487110 1M0316M21
A1978307 R273.R.EC
BB583147 BB583147
AA807648 nx08c03.s
BE516407 WHE610.H0
Z37222 ATTS3965.AC
AA634103 n109h06.s
AV154082 AV154082
AZ728927 RPCI-24-1
AM214673 up01a03.y
AO388893 RPCI11-14
BE810749 MR4-LT007
AQ027837 CIT-HSP-2
BB118069 BB118069
BE528278 W71095STM
BG487311 d195f04.y
BB601852 BB601852
AZ020088 RPCI-23-3
R16206 y4d9b12.s2
A1441353 sas5f06.y
AA825331 oe63b06.s
BB309167 BB309167
BB412379 BB412379
AA096629 mo03d10.r
AT002342 AT002342
AZ587088 1M0394A06
BB176196 BB176196
AU029833 AU029833
AM853559 RC3-CT005
AZ500749 1M0339J15
AV076081 AV076081
AV077408 AV077408
BB319896 BB319896
BB572232 BB572232
AM322270 uc05f09.x
BB603068 BB603068
W63249 T9ESTzy8h10
BG014672 QV2-GN020
AA319807 EST22283
AU293702 UT-H-B12-
AZ625139 1M0464M14
BB228806 BB228806
Z21331 HSAADXX.T
AZ419333 1M0195J15
BB117978 BB117978
AO922536 RPCI-23-2
BB366546 BB366546
AZ588220 1M0396120
AM230121 up25f12.y
BB262861 BB262861
BA452180 BB452180

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 894)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapds-rt@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://imgc.llnl.gov>
Plate: LICM1435 row: 0 column: 20
High quality sequence stop: 759.

FEATURES

Location/Qualifiers

1..894
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4652011"
/clone_lib="NIH-MGC_21"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 205 a 248 c 250 g 190 t 1 others
ORIGIN

Query Match 2.2%; Score 21; DB 154; Length 894;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 932 acatcacgaagggtgaagc 952
|||||
Db 825 ACATCAGAGAGGGGTGAAGC 845

RESULT 2
AA645779/c 222 bp mRNA EST 28-OCT-1997
LOCUS vnl0h04.r1 Stratigene mouse Tcell 937311 Mus musculus CDNA clone
DEFINITION IMAGE:1020823 5', mRNA sequence.
ACCESSION AA645779
VERSION AA645779.1 GI:2572208
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 222)
REFERENCE Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisler, S., Kucab, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@imgc.llnl.gov) for further information.
MGI:571599
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 220.

FEATURES

Location/Qualifiers

1..222
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1020823"
/clone_lib="Stratigene mouse Tcell 937311"
/tissue_type="Tcell"
/dev_stage="M30 CD4+ cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: blood; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. M30 CD4+ cells. Average insert size 1.0 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG
3'-3' adaptor sequence: 5' CTCGACGTTTCTTTTCTTTT 3'."
BASE COUNT 72 a 45 c 53 g 52 t
ORIGIN

Query Match 2.0%; Score 20; DB 10; Length 222;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 781 caactctgaatctggatt 800
|||||
Db 21 CAACCTCTGATCTGGCTT 2

RESULT 3
AM268708/c 414 bp mRNA EST 03-JAN-2000
LOCUS xv35b05.x1 Soares_NFL.T.GBC.S1 Homo sapiens CDNA clone
DEFINITION IMAGE:2815089 3', mRNA sequence.
ACCESSION AM268708
VERSION AM268708.1 GI:6655738
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 414)
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapds-rt@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@imgc.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 408.

FEATURES

Location/Qualifiers

1..414
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2815089"
/clone_lib="Soares_NFL.T.GBC.S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung Nhlh19W, testis NHT, and B-cell
NCI-CGAP-GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT 87 a 106 c 80 g 141 t
ORIGIN

Query Match 2.0%; Score 20; DB 113; Length 414;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 813 gacagagctgcagcagcc 832
|||||
Db 158 GACAGAGCTGCAGCAGCC 139

RESULT 4
LOCUS A1769658 438 bp mRNA EST 21-DEC-1999
DEFINITION wj2502.x1 NCI-CGAP_K1d12 Homo sapiens cDNA clone IMAGE:2403843 3',
mRNA sequence.
ACCESSION A1769658
VERSION A1769658.1 GI:5236167
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmett-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html
Insert Length: 536 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers

FEATURES
source 1..438
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2403843"
/clone_lib="NCI-CGAP_K1d12"
/issue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: p7T73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI-CGAP_K1d5 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(clonoids 1323912-1325831, 1471368-1472903 and
1492104-1493255). Subtraction by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT 93 a 110 c 89 g 146 t
ORIGIN

Query Match 2.0%; Score 20; DB 24; Length 438;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 813 gacagagctgcagcagcc 832
|||||
Db 159 GACAGAGCTGCAGCAGCC 140

RESULT 5
LOCUS A1028162/c 456 bp mRNA EST 27-AUG-1998
DEFINITION ov90402.x1 Soares-testis_NHT Homo sapiens cDNA clone IMAGE:1644554
3', mRNA sequence.
ACCESSION A1028162
VERSION A1028162.1 GI:3245471
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html
Insert Length: 835 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 359.
Location/Qualifiers

FEATURES
source 1..456
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1644554"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: p7T73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5'
TGTACCAATCTGACAGCGGCGCCCAATTTTCTTTTCTTTT 3']
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p7T73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 102 a 110 c 93 g 151 t
ORIGIN

Query Match 2.0%; Score 20; DB 15; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 813 gacagagctgcagcagcc 832
|||||
Db 160 GACAGAGCTGCAGCAGCC 141

RESULT 6
LOCUS AM269670/c 483 bp mRNA EST 03-JAN-2000
DEFINITION xv36d11.x1 Soares_NFL_T-GBC_S1 Homo sapiens cDNA clone
IMAGE:2815221 3', mRNA sequence.
ACCESSION AM269670
VERSION AM269670.1 GI:6656700
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html
Insert Length: 835 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 359.
Location/Qualifiers

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbrr@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 471.
Location/Qualifiers

FEATURES
source
1. 483
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2815221"
/clone_1lb="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: p773D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI-CGAP-GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT 105 a 121 c 100 g 157 t
ORIGIN

Query Match 2.0%; Score 20; DB 113; Length 483;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 813 gacagagctgcagcagacc 832
|||||
Db 158 GACAGAGCTGCAGCAGACC 139

RESULT 7
A1594177 490 bp mRNA EST 21-APR-1999
LOCUS vni0h04.y1 Stratagene mouse Tcell 937311 Mus musculus cDNA clone
DEFINITION IMAGE:1020823 5', mRNA sequence.
ACCESSION A1594177
VERSION A1594177.1 GI:4603225
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
1 (bases 1 to 490)
Marrar,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,T., Person
B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritzer
E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The Mashu-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marra M/Mashu-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:571599
This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the

correct orientation)
Seq primer: -40RP from Gibco
High quality sequence stop: 397.
Location/Qualifiers

FEATURES
source
1. 490
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1020823"
/clone_1lb="Stratagene mouse Tcell 937311"
/tissue_type="Tcell"
/dev_stage="M30 CD4+ cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: blood; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. M30 CD4+ cells. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCGACGAG
3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3'."

BASE COUNT 174 a 94 c 117 g 105 t
ORIGIN

Query Match 2.0%; Score 20; DB 22; Length 490;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 781 caactctgaatctggatt 800
|||||
Db 22 CAACCTCTGCACTCGGCTT 3

RESULT 8
BE893506 917 bp mRNA EST 20-OCT-2000
LOCUS BE893506/c 601438060F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3923007 5',
DEFINITION mRNA sequence.
ACCESSION BE893506
VERSION BE893506.1 GI:10354930
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 917)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbrr@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Place: LLM9758 row: h column: 16
High quality sequence stop: 607.
Location/Qualifiers

FEATURES
source
1. 917
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3923007"
/clone_1lb="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORE6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 226 a 219 c 319 g 153 t
ORIGIN

Query Match 2.0%; Score 20; DB 141; Length 917;
 Best Local Similarity 100.0%; Pred. No. 8.4;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 671 accagtcgagggcaggtt 690
 ||||||||||||||||
 Db 917 ACCAGTCGAGGGCAGGCTT 898

RESULT 9
 A0511509/c
 LOCUS A0511509 468 bp DNA GSS 05-MAY-1999
 DEFINITION HS-5074.B1.C08.SP6E RPCI-11 Human Male BAC Library Homo sapiens
 genomic clone Plate=650 COL=15 Row=F, DNA sequence.
 ACCESSION A0511509
 VERSION A0511509.1 GI:4744119
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 468)
 AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieterdejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
 or from Research Genetics (info@resgen.com). BAC end Web Server:
 http://www.htsc.washington.edu
 Plate: 650 row: F column: 15
 Seq primer: SP6
 Class: BAC ends
 High quality sequence stop: 468.

FEATURES source

1. 468
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="Plate=650 COL=15 Row=F"
 /clone_1id="RPCI-11 Human Male BAC Library"
 /sex="male"
 /note="Vector: pBAC3.6; Site.1: EcoRI; Site.2: EcoRI;
 Male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI Methylase. Size selected DNA was cloned into the
 pBAC3.6 vector at EcoRI sites"
 BASE COUNT 143 a 92 c 92 g 135 t 6 others
 ORIGIN

Query Match 1.9%; Score 19; DB 229; Length 468;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 580 agctcgtgcgcataatg 598
 ||||||||||||||||
 Db 39 AGCTCGTGGCGATATATG 21

RESULT 10
 A0337527

LOCUS A0337527 523 bp DNA GSS 06-MAR-1999
 DEFINITION HS-5017.A2.F02.T7 RPCI11 Human Male BAC Library Homo sapiens
 genomic clone Plate=593 COL=4 Row=K, DNA sequence.
 ACCESSION A0337527
 VERSION A0337527.1 GI:4141739
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 523)
 AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones may be purchased from Research Genetics (info@resgen.com).
 BAC end Web Server: http://www.htsc.washington.edu
 Plate: 593 row: K column: 4
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 523.

FEATURES source

1. 523
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="Plate=593 COL=4 Row=K"
 /clone_1id="RPCI11 Human Male BAC Library"
 /sex="Male"
 /cell_type="lymphocytes"
 /note="Vector: pBAC3.6; RPCI11 Human Male BAC Library"
 BASE COUNT 183 a 151 c 125 g 58 t 6 others
 ORIGIN

Query Match 1.9%; Score 19; DB 227; Length 523;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 347 ggattaatgtaccacaag 365
 ||||||||||||||||
 Db 151 GGATTAAATGTCACCCAAAG 169

RESULT 11
 A2497878/c
 LOCUS A2497878 656 bp DNA GSS 05-OCT-2000
 DEFINITION IM0334G24R Mouse 10kb plasmid UUC1M library Mus musculus genomic
 clone UUC1M0334G24 R, DNA sequence.
 ACCESSION A2497878
 VERSION A2497878.1 GI:10675204
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 656)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
 and Wright,D., Weiss,R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)

Query Match 1.9%; Score 19; DB 221; Length 713;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 aggaagaacagcagccag 124
|||||
Db 650 AGGAGAGAACAGCAGCCAG 632

RESULT 14
AI325737 747 bp mRNA EST 23-DEC-1998
LOCUS mm95003.y1 Striatagene mouse heart (#937316) Mus musculus cDNA clone
DEFINITION IMAGE:536165 5' similar to gb:X16869 ELCONGATION FACTOR 1-ALPHA 1
(HUMAN); gb:X13661 Mouse mRNA for elongation factor 1-alpha (MOUSE
); mRNA sequence.
ACCESSION AI325737
VERSION AI325737.1 GI:4060166
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 747)
REFERENCE Maria,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
JOURNAL Contact: Maria M/Mouse EST Project
COMMENT WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:353101
This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the
correct orientation)
Seg primer: -40RP from Gibco
High quality sequence stop: 114.
Location/Qualifiers
1. 747
/organism="Mus musculus"
/strain="NIH/Swiss"
/db_xref="taxon:10090"
/clone="IMAGE:536165"
/clone_lib="Striatagene mouse heart (#937316)"
/sex="pooled"
/tissue_type="heart"
/dev_stage="13 day embryos"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: heart; Vector: pBluescript SK-; Site:1:
EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. 93 pooled NIH/Swiss 13 day embryo hearts.
Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5',
adaptor sequence: 5' GAATTCGGCAGCG 3' -3' adaptor
sequence: 5' CTCGAGCTTTTCTTTTCTTTTCTT 3'."

BASE COUNT 208 a 182 c 148 g 208 t 1 others
ORIGIN

Query Match 1.9%; Score 19; DB 18; Length 747;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 364 aggaagctgcagcagcactg 382
|||||

Db 573 AGGAGCTGCTCAGCAGCTG 591

RESULT 15
A2668255/c 842 bp DNA GSS 14-DEC-2000
LOCUS ENTKX52TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic, DNA sequence.
ACCESSION A2668255
VERSION A2668255.1 GI:11805401
KEYWORDS GSS.
SOURCE Entamoeba histolytica.
ORGANISM Entamoeba histolytica.
REFERENCE 1 (bases 1 to 842)
AUTHORS Loftus,B., Van Aken,S. and Fraser,C.
TITLE Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library
Unpublished (2000)
JOURNAL Contact: Brendan J Loftus
COMMENT Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: bjloftus@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 24
High quality sequence stop: 731.
Location/Qualifiers
1. 842
/organism="Entamoeba histolytica"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOSt; Site:1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
Whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."

BASE COUNT 356 a 92 c 112 g 282 t
ORIGIN

Query Match 1.9%; Score 19; DB 247; Length 842;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 776 atgaacacttcgatct 794
|||||
Db 800 ATGACAACTCTGTAATCT 782

RESULT 16
CNS03YOE 891 bp DNA GSS 18-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
DEFINITION 060J20 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL266639.1 GI:7988418
VERSION AL266639.1
KEYWORDS GSS; genome survey sequence.

| SOURCE | ORGANISM | TITLE | REFERENCE | AUTHORS |
|-----------------------|--|---------------------------------|--|---------|
| FEATURES | source | JOURNAL REFERENCE AUTHORS | Genoscope. Direct Submission Submitted (12-APR-2000) to the EMBL/GenBank/DBD databases This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon . Location/Qualifiers 1..891 /organism="Tetraodon nigroviridis" /db_xref="taxon:99883" /clone="068J20" /clone_1id="G" /note="Genoscope sequence ID : CDBG068DE10SP1-end : PUC-Or1" | |
| BASE COUNT | 254 a 196 c 213 g 216 t 12 others | | | |
| ORIGIN | | | | |
| Query Match | 1.9%, Score 19; DB 221; Length 891; | | | |
| Best Local Similarity | 100.0%; Pred. No. 31; | | | |
| Matches | 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | |
| OR | 231 cccctgcattcccccctgtt 249 | | | |
| DB | 143 CCCTGCATCACCCTCCTGT 161 | | | |
| RESULT 17 | | | | |
| LOCUS | BB043716 | 235 bp | mRNA | EST |
| DEFINITION | BB043716 RIKEN full-length enriched, 13 days embryo male testis mus musculus cDNA clone 6030477B18 3' similar to AL133589 Homo sapiens | | | |
| ACCESSION | BB043716 | | | |
| VERSION | BB043716.1 | GI:8450102 | | |
| KEYWORDS | EST. | | | |
| SOURCE | house mouse. | | | |
| ORGANISM | Mus musculus | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| AUTHORS | 1 (bases 1 to 235) Konno,H., Alizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Horii,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Iizawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kituchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata Y., Shigemoto,Y., Shinagawa,A., Shitaki,T., Sogabe,Y., Sugihara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toya T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamataka,I. | | | |

| | |
|---|--|
| TITLE | JOURNAL |
| COMMENT | |
| <p>'Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,N. and Hayashizaki,Y. RIKEN Mouse ESTs (Kono,H., et al.) Unpublished (2000) Contact: Yoshihide Hayashizaki Genome Exploration Research Group, Life Science Tsukuba Center, Genome Science Laboratory The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel.: +81-298-36-9013 Fax: +81-298-36-9098 Email: genome-res@rtc.riken.go.jp/ URL:http://genome.rtc.riken.go.jp/ Carninci,P., Nishiyama,Y., Westover,A., Itch,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Thermotabilization and thermoactivation of thermostable enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998) Itch,M., Kitsuina,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y. Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999) Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999) Please visit our web site (http://genome.rtc.riken.go.jp) for further details.</p> | |
| FEATURES | SOURCE |
| <p>Location/Qualifiers 1..235 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="6030477B18" /clone_lib="RIKEN full-length enriched, 13 days embryo male testis" /sex="male" /tissue_type="testis" /dev_stage="13 days embryo" /lab_host="DH10B" /note="Site.1: SalI; Site.2: BamHI. CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGGACGAGAATTCGCAGTAAATAAATTATTAATCCCCCCCCCCC 3'] CDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. CDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adaptor of sequence [5' GAGGACGAGAATTCGCAGTAAATAAATTATTAATCCCCCCCCCCC 3'] CDNA was cloned into the XhoI and BamH sites. Vector: a modified Bluescript KS(+) after bulk excision from Lambda phage I. Cloning sites, 5' end: SalI; 3' end: BamHI."</p> | |
| BASE COUNT | 60 a 69 c 41 g 65 t |
| ORIGIN | |
| <p>Query Match 1.8%; Score 18; DB 124; Length 235; Best Local Similarity 100.0%; Pred. No. 1.1e+02; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;</p> | |
| Ox | 111 agaacagcagccagcgcca 128 Db 1 AGAACAGCAGCGCCAGGCCA 18 |
| RESULT 18 | AV292150/c |

LOCUS AV292150 241 bp mRNA EST 09-NOV-1999
 DEFINITION AV292150 RIKEN full-length enriched, 6 days neonate head Mus musculus cDNA clone 5430416G10 3', mRNA sequence.
 ACCESSION AV292150
 VERSION AV292150.1 GI:6306181
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Mammalia; Eutheria; Rodentia; Sclerognathia; Muridae; Murine; Mus.
 1 (bases 1 to 241)
 REFERENCE Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, K., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shitaki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateo, M., Tomioka, N., Tsunoda, Y., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
 RIKEN Mouse ESTs (Kono, H., et al. 1999)
 TITLE Unpublished (1999)
 JOURNAL Contact: Yoshihide Hayashizaki
 COMMENT Genome Exploration Research Group, Life Science Tsukuba Center, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: +81-298-36-9013
 Fax: +81-298-36-9098
 Email: genome-res@rtrc.riken.go.jp,
 URL: http://genome.rtrc.riken.go.jp/
 Sasaki, N., Izawa, M., Watanabe, M., Okazaki, Y., and Hayashizaki, Y., Matsura, S., Carninci, P., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y., 1.
 Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
 Itoh, M., Katsunari, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
 Please visit our web site (http://genome.rtrc.riken.go.jp) for further details.
 FEATURES
 source location/Qualifiers
 1. 241
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="5430416G10"
 /clone_lib="RIKEN full-length enriched, 6 days neonate head"
 /sex="mixed"
 /tissue_type="head"
 /dev_stage="6 days neonate"
 /lab_host="DH10B"
 /note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
 GAGAGAGAGAGATCCAGAGCTTTTCTTTTCTTTTCTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of

sequence [5' GAGAGAGAGATTCGAGTTAATTAATTAATTCACCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified plasmid pBluescript KS(+) after bulk excision from lambda FLIC I. Cloning sites, 5' end: SalI; 3' end: BamHI"
 BASE COUNT 70 a 37 c 33 g 101 t
 ORIGIN
 Query Match 1.8%; Score 18; DB 28; Length 241;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Yy 774 ccaatgaacactctgaa 791
 ||||||||||||||||
 Db 124 CCATGACACTCTCGAA 107
 RESULT 19
 AV526879/c
 LOCUS AV526879 244 bp mRNA EST 01-SEP-2000
 DEFINITION AV526879 Arabidopsis thaliana aboveground organs two to six-week old Arabidopsis thaliana cDNA clone AP23103R 5', mRNA sequence.
 ACCESSION AV526879
 VERSION AV526879.1 GI:8686407
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 244)
 REFERENCE Asamizu, E., Nakamura, Y., Sato, S., and Tabata, S.
 A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
 DNA Res. 7, 175-180 (2000)
 JOURNAL Contact: Erika Asamizu
 MEDLINE The First Laboratory for Plant Gene Research
 COMMENT Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
 FEATURES
 source location/Qualifiers
 1. 244
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="AP23103R"
 /clone_lib="Arabidopsis thaliana aboveground organs two to six-week old"
 /tissue_type="aboveground organs"
 /dev_stage="two to six-week old"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"
 BASE COUNT 74 a 67 c 44 g 59 t
 ORIGIN
 Query Match 1.8%; Score 18; DB 31; Length 244;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Yy 930 caatcaatgaaggggt 947
 ||||||||||||||||
 Db 140 CACATCAGCAGAGGGGT 123
 RESULT 20
 AA358477/c
 LOCUS AA358477 261 bp mRNA EST 21-APR-1997
 DEFINITION EST67411 Fetal Lung III Homo sapiens cDNA 5' end, mRNA sequence.
 ACCESSION AA358477

```

VERSION AA358477.1 GI:2010845
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 261)
JOURNAL Adams,M.D., Soares,M.B., Kerlavage,A.R., Fields,C. and Venter,J.C.
MEDLINE Rapid cDNA sequencing (expressed sequence tags) from a
COMMENT directionally cloned human infant brain cDNA library
94004965
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M3 Reverse.

FEATURES
Location/Qualifiers
1..261
/organism="Homo sapiens"
/db_xref="ATCC (lnhost):162671"
/db_xref="taxon:9606"
/clone.lib="Fetal Lung IIT"
/dev_stage="fetus"
/note="organ:lung; Vector: pBluescript SK-; Site_1: EcoRI
; Site_2: XhoI"

BASE COUNT 47 a 77 c 85 g 48 t 4 others

ORIGIN

Query Match 1.8%; Score 18; DB 6; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 gtagctgcgtcctgcgtg 61
|||||
Db 18 gTACGTGCTCTCTCGCTG 1

RESULT 21
AV527317 264 bp mRNA EST 01-SEP-2000
LOCUS AV527317 Arabidopsis thaliana aboveground organs two to six-week
DEFINITION Old Arabidopsis thaliana cDNA clone AP234b05R 5', mRNA sequence.
ACCESSION AV527317
VERSION AV527317.1 GI:8686845
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
TITLE 1 (bases 1 to 264)
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
REFERENCE A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL DNA Res. 7, 175-180 (2000)
MEDLINE 20363093
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yama 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1..264
/organism="Arabidopsis thaliana"

FEATURES
source

```

```

/strain="Columbia"
/db_xref="taxon:3702"
/clone.lib="AP234b05R"
/clone.lib="Arabidopsis thaliana aboveground organs two to
six-week old"
/tissue_type="aboveground organs"
/dev_stage="two to six-week old"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 85 a 71 c 49 g 59 t

ORIGIN

Query Match 1.8%; Score 18; DB 31; Length 264;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 930 caacatcacgaaggaggt 947
|||||
Db 152 CAACATCACGACGAGGGGCT 135

RESULT 22
Z29054/c 294 bp mRNA EST 16-DEC-1993
LOCUS Z29054 Arabidopsis thaliana cDNA clone OA067 5' similar to
DEFINITION complex II type III chlorophyll a/b binding prot, mRNA sequence.
ACCESSION Z29054
VERSION Z29054.1 GI:435254
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
TITLE 1 (bases 1 to 294)
AUTHORS CNRS.
REFERENCE The Arabidopsis thaliana transcribed genome: the GDR cDNA program
JOURNAL Unpublished (1996)
COMMENT Contact: Krivitzky M., Bonnet R., Jean-Jacques I., Kreis M.,
Lecharny A.
Laboratoire de Biologie du Développement des Plantes
Université de Paris Sud
Bat. 450, 91405 Orsay Cedex, France
Email: alainedp.ups.circe.fr.
Location/Qualifiers
1..294
/organism="Arabidopsis thaliana"
/strain="ecotype Columbia"
/db_xref="taxon:3702"
/clone.lib="OA067"
/clone.lib="Ora-A"
/tissue_type="green shoots"
/note="Vector: lambda ZAPII"

BASE COUNT 76 a 78 c 65 g 75 t

ORIGIN

Query Match 1.8%; Score 18; DB 191; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 930 caacatcacgaaggaggt 947
|||||
Db 121 CAACATCACGACGAGGGGCT 104

RESULT 23
AV208201 301 bp mRNA EST 30-OCT-1999
LOCUS AV208201 RIKEN full-length enriched, adult male testis Mus musculus
DEFINITION cDNA clone 1700106B02 3' similar to AF065366 Mus musculus Jun
coactivator Jabi (Jabi) mRNA, mRNA sequence.

```


Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 318)
 CNRS.
 The Arabidopsis thaliana transcribed genome: the GDR cDNA program
 Unpublished (1996)
 Contact: Berthomieu P., Guerrier D., Giraudat J.
 Genetique Moleculaire d'Arabidopsis
 ISV - UPR40, CNRS
 Avenue de la Terrasse, 1198 Gif-sur-Yvette Cedex, France
 Email: Giraudat@cnrs-gif.fr
 CDS 6. .>318 /product="CHLOROPHYLL A-B BINDING PROTEIN PRECURSOR".
 Location/Qualifiers
 1. 318
 /organism="Arabidopsis thaliana"
 /strain="ecotype Columbia"
 /db_xref="taxon:3702"
 /clone_1ib="AY058"
 /clone_1ib="AC16H"
 /tissue_type="cell suspension culture"
 /dev_stage="cycling cells"
 /note="Vector: lambda ZAP11"
 /note="Vector: lambda ZAP11"

BASE COUNT 80 a 86 c 70 g 82 t
 ORIGIN

Query Match 1.8%; Score 18; DB 191; Length 318;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 930 caacatcaggaagggt 947
 ||||||||||||||||
 Db 102 CACATCAGCAGAGGGGT 85

RESULT 26
 AM896020 327 bp mRNA EST 24-MAY-2000
 LOCUS
 DEFINITION QV4-NN0039-110500-204-a03 NN0039 Homo sapiens cDNA, mRNA sequence.
 AM896020
 VERSION AM896020.1 GI:8060225
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 327)
 Dias Neto, E., Garcia Correia, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Balda, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?fl=612-QV4-NN0039-110
 500-204-a03&ct3=2000-05-11&ct4=1)
 High quality sequence start: 9
 High quality sequence stop: 132.

FEATURES Location/Qualifiers
 source 1. 327
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1ib="NN0039"
 /dev_stage="Adult"
 /note="Organ: nervous, normal; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

BASE COUNT 107 a 65 c 98 g 57 t
 ORIGIN

Query Match 1.8%; Score 18; DB 122; Length 327;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 544 acagagagaagaagcc 561
 ||||||||||||||||
 Db 140 ACAGGAGAGAGAAAGGCC 157

RESULT 27
 T21227/c 331 bp mRNA EST 06-NOV-1997
 LOCUS
 DEFINITION 3235 Lambda-PRL2 Arabidopsis thaliana cDNA clone 93E117, mRNA
 sequence.
 T21227
 VERSION T21227.1 GI:2596344
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 331)
 Newman, T., deBurlin, F.J., Green, P., Keegstra, K., Kende, H., McIntosh,
 L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel,
 E. and Somerville, C.
 Genes galore: a summary of methods for accessing results from
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 Plant Physiol. 106, 1241-1255 (1994)
 95148729
 On Nov 6, 1997 this sequence version replaced gi:931743.
 Contact: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PR, Michigan State University, Plant Biology Bldg., E.
 Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313tc@ldm.cl.msu.edu
 Seq primer: T7 dye primer.
 Location/Qualifiers
 1. 331
 /organism="Arabidopsis thaliana"
 /strain="var columbia"
 /db_xref="taxon:3702"
 /clone_1ib="93E117"
 /clone_1ib="Lambda-PRL2"
 /note="Vector: lambda Zip-lox; Site_1: SalI; Site_2: NotI;
 Lambda PRL2 is a cDNA library derived from equal
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7
 day germinated etiolated seedlings; 2) tissue culture
 grown roots; 3) staged plants half with 24 hour light
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
 same plants as 3 but aerial tissue (stems, flowers and
 siliques). The vector is BRV's lambda Zip-lox. The cDNA
 inserts were directionally cloned with Sal-Not arms using

MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 cDNA Library Preparation: M.B. Soares Lab Clone distribution:
 clones will be available through Research Genetics (www.resgen.com)
 This clone is also available through the I.M.A.G.E. Consortium at
 LLNL (info@image.llnl.gov). IMAGE ID= 1792965
 Seq primer: M3 forward.

FEATURES
 source
 1. 340
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-B70-px-d-11-0-UI"
 /clone_1lb="UI-R-B70"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: PT73D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I. Site_2: Eco RI. This library
 (UI-R-B70) consists of a mixture of individually tagged
 normalized libraries constructed from rat hippocampus,
 thalamus, mid-brain, medulla, corpus striatum, cerebral
 cortex and testis. The tag used to identify the source
 tissue is a string of 3-6 nucleotides present between the
 Not I site and the oligo-dT track which allows
 identification of the library of origin of a clone within
 the mixture. This library was then subtracted using a
 driver consisting of a mixture of all clones from UI-R-A0,
 UI-R-A1, UI-R-E0, UI-R-E1, UI-R-C0, UI-R-C1, UI-R-C2 and
 UI-R-C2p."

BASE COUNT 82 a 82 c 78 g 98 t
 ORIGIN

Query Match 1.8%; Score 18; DB 150; Length 340;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 812 tgacagagctgcagcaga 829
 ||||||||||||||||
 Db 142 TGACAGAGCTGCAGCAGA 159

RESULT 31
 BE520970/c 342 bp mRNA EST 19-MAR-2001
 LOCUS M16A2STM Arabidopsis developing seed Arabidopsis thaliana CDNA
 DEFINITION clone M16A2 5', mRNA sequence.
 ACCESSION BE520970
 VERSION BE520970.1 GI:9778948
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 342)
 White,J.A., Todd,J., Newman,T., Focks,N., Gilke,T., Martinez de
 Illarduya,O., Jaworski,J.G., Ohlrogge,J. and Banning,C.
 A new set of Arabidopsis expressed sequence tags from developing
 seeds. The metabolic pathway from carbohydrates to seed oil
 Plant Physiol. 124 (4), 1582-1594 (2000)
 20367808
 CONTACT: Banning, C
 Dept. of Biochemistry & Molecular Biology
 Michigan State University
 224 Biochemistry, Michigan State University, East Lansing, MI 48824
 , USA

Tel: 517 355 1609
 Fax: 517 353 9334
 Email: banning@msu.edu
 Michigan State University DNA Sequencing Facility Arabidopsis
 Biological Resource Center, The Ohio State University, 309 Botany &
 Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
 6142920603 TEL: 6142929371.
 Location/Qualifiers

FEATURES
 source
 1. 342
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="M16A2"
 /clone_1lb="Arabidopsis developing seed"
 /tissue_type="seed"
 /dev_stage="5-13 days after flowering"
 /lab_host="E.coli"
 /note="Organ: Developing seed; Vector: pBluescript SK-;
 Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 82 a 90 c 80 g 90 t
 ORIGIN

Query Match 1.8%; Score 18; DB 136; Length 342;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 930 caacatcacgaaggaggt 947
 ||||||||||||||||
 Db 108 CAACATCAGCAGAGCGGCT 91

RESULT 32
 AV527474 365 bp mRNA EST 01-SEP-2000
 LOCUS AV527474/c Old Arabidopsis thaliana aboveground organs two to six-week
 DEFINITION old Arabidopsis thaliana cDNA clone AP237e11R 5', mRNA sequence.
 ACCESSION AV527474
 VERSION AV527474.1 GI:8687002
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 365)
 Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
 A large scale analysis of cDNA in Arabidopsis thaliana: Generation
 of 12,028 non-redundant expressed sequence tags from normalized and
 size-selected cDNA libraries
 DNA Res. 7, 175-180 (2000)
 20363093
 CONTACT: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/
 Location/Qualifiers

FEATURES
 source
 1. 365
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="AP237e11R"
 /clone_1lb="Arabidopsis thaliana aboveground organs two to
 six-week old"
 /tissue_type="aboveground organs"
 /dev_stage="two to six-week old"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
 XhoI"

BASE COUNT 103 a 96 c 81 g 85 t
 ORIGIN

Query Match 1.8%; Score 18; DB 31; Length 365;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 930 caacatcagaagggt 947
|||||
Db 142 CAACATCAGGAGGGGT 125

RESULT 33
T42951/c 368 bp mRNA EST 06-NOV-1997
LOCUS 6214 Lambda-PRL2 Arabidopsis thaliana cDNA clone 118112T7, mRNA
DEFINITION sequence.
T42951
ACCESSION T42951 GI:2597542
VERSION T42951.1
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 368)
AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel
,E. and Somerville,C.
TITLE Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
JOURNAL Plant Physiol. 106, 1241-1255 (1994)
MEDLINE 95148729
COMMENT On Nov 6, 1997 this sequence version replaced gi:947584.
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313c@clim.cl.msu.edu
Seq primer: 77 dye primer.
Location/Qualifiers
1. 368
/organism="Arabidopsis thaliana"
/strain="var Columbia"
/db_xref="taxon:3702"
/clone="118112T7"
/clone_id="Lambda-PRL2"
/note="Vector: lambda zip-lox; Site_1: Sal; Site_2: Not;
Lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques). The vector is BRU's lambda zip-lox. The cDNA
inserts were directionally cloned with Sal-I Not arms using
oligo dT primed cDNA."

BASE COUNT 91 a 86 c 83 g 94 t 14 others
ORIGIN

Query Match 1.8%; Score 18; DB 189; Length 368;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 930 caacatcagaagggt 947
|||||
Db 139 CAACATCAGGAGGGGT 122

RESULT 34
T47675/c 370 bp mRNA EST 17-JAN-1995
LOCUS

DEFINITION ATTS4531 Strasbourg-A Arabidopsis thaliana cDNA clone FAFJ47, mRNA
sequence.
ACCESSION 247675
VERSION 247675.1 GI:623714
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 370)
AUTHORS CNRS.
TITLE The Arabidopsis thaliana transcribed genome: the GDR cDNA program
JOURNAL Unpublished (1996)
COMMENT Contact: Parmentier Y., Criqui M.C., Durr A., Fleck J.
Fleck Jacqueline / U626
Biologie Moleculaire des Plantes - CNRS
12 Rue du General Zimmer, 67084 Strasbourg Cedex, France
Email: ARABANK@MEDOC.U-STRASBG.FR.
Location/Qualifiers
1. 370
/organism="Arabidopsis thaliana"
/strain="ecotype Columbia"
/db_xref="taxon:3702"
/clone="FAFJ47"
/clone_id="Strasbourg-A"
/note="Vector: Lambda ZAPIT; tissue-sliced leaves of
A.thaliana ecotype Columbia; clone library-Strasbourg-A;
Cloning vector: Lambda ZAPIT; Physiological condition:
leaves strips incubated 2/3/4 days in liquid culture
medium."

BASE COUNT 87 a 96 c 85 g 100 t 2 others
ORIGIN

Query Match 1.8%; Score 18; DB 191; Length 370;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 930 caacatcagaagggt 947
|||||
Db 96 CAACATCAGGAGGGGT 79

RESULT 35
AV430174/c 376 bp mRNA EST 23-AUG-2000
LOCUS AV430174 Porphyra yezoensis TU-1 Porphyra yezoensis cDNA clone
PL014e12_r 5', mRNA sequence.
DEFINITION AV430174
ACCESSION AV430174 GI:8585399
KEYWORDS EST.
KEYWORDS Porphyra yezoensis.
SOURCE Porphyra yezoensis
ORGANISM Eukaryota: Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyra.
REFERENCE 1 (bases 1 to 376)
AUTHORS Nikaido,I., Asamizu,E., Nakajima,M., Nakamura,Y., Saga,N. and
Tabata,S.
TITLE Generation of 10,154 expressed sequence tags from a leafy
gametophyte of a marine red alga, Porphyra yezoensis
JOURNAL DNA Res. 7, 223-227 (2000)
MEDLINE 20363100
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizukazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.
Location/Qualifiers
1. 376
/organism="Porphyra yezoensis"
/strain="TU-1"
/db_xref="taxon:2788"

```

BASE COUNT      40 a      158 c      121 g      57 t
ORIGIN
/clone="PL014e12_r"
/clone_lib="Porphyra yezoensis TU-1"
/Note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

Query Match      1.8%; Score 18; DB 30; Length 376;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 596 atcgacagccggcgccg 613
|||||
Db 220 ATGCACACGCCCGCGCC 203

RESULT 36
LOCUS AV526330 391 bp mRNA EST 01-SEP-2000
DEFINITION AV526330 Arabidopsis thaliana aboveground organs two to six-week
ACCESSION AV526330 Old Arabidopsis thaliana cDNA clone AP210d10R 5', mRNA sequence.
VERSION AV526330.1 GI:8685858
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 391)
Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)

JOURNAL 20363093
MEDLINE Contact: Erika Asamizu
COMMENT The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp. URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1..391
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="AP210d10R"
/clone_lib="Arabidopsis thaliana aboveground organs two to
six-week Old"
/issue_type="aboveground organs"
/dev_stage="two to six-week Old"
/Note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT      107 a      100 c      87 g      97 t
ORIGIN
/clone="PL014e12_r"
/clone_lib="Porphyra yezoensis TU-1"
/Note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

Query Match      1.8%; Score 18; DB 31; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 930 caacatcaggaagggt 947
|||||
Db 134 CAACATCAGCAGGCGGT 117

RESULT 37
LOCUS N65689 391 bp mRNA EST 05-JAN-1998
DEFINITION N65689 20729 Lambda-PRL2 Arabidopsis thaliana cDNA clone 229K107, mRNA
sequence.
ACCESSION N65689

```

```

VERSION N65689.1 GI:1217315
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 391)
Newman, T., deBrujin, F.J., Green, P., Keegstra, K., Kende, H., McIntosh
L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel
E., and Somerville, C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)

TITLE Contact: Thomas Newman
JOURNAL MSU-DOE Plant Research Laboratory
MEDLINE Michigan State University
95148729 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313cpeidm.cl.msu.edu

FEATURES
source
1..391
/organism="Arabidopsis thaliana"
/strain="var columbica"
/db_xref="taxon:3702"
/clone="229K107"
/clone_lib="Lambda zip-lox"
/Note="Vector: lambda zip-lox; Site_1: Sal; Site_2: Not;
Lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark - rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques). The vector is BRL's lambda zip-lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA."

BASE COUNT      100 a      86 c      90 g      100 t      15 others
ORIGIN
/clone="PL014e12_r"
/clone_lib="Porphyra yezoensis TU-1"
/Note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

Query Match      1.8%; Score 18; DB 159; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 930 caacatcaggaagggt 947
|||||
Db 147 CAACATCAGCAGGCGGT 130

RESULT 38
LOCUS D78804 392 bp mRNA EST 09-FEB-1996
DEFINITION HUM516H08B Human Placenta polyA+ (T Fujiiwara) Homo sapiens cDNA
clone GEN-516H08 5', mRNA sequence.
ACCESSION D78804
VERSION D78804.1 GI:1180677
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 392)
Fujiiwara, T., Hirano, H., Katagiri, T., Kawai, A., Kuga, Y., Nagata, M.,
Okuno, S., Ozaki, K., Shimizu, F., Shimada, Y., Shimomiyu, H., Takaichi
A., Takeda, S., Watanabe, T., Takahashi, E., Hirai, Y., Maekawa, H.,
Shin, S. and Nakamura, Y.
Fujiiwara et al. (1995)
unpublished (1995)

```

COMMENT

Contact: Tsutomu Fujiwara
Otsuka GEN Research Institute
Otsuka Pharmaceutical Co. Ltd
463-10 Kagasuno Kawanchi-cho, Tokushima, Tokushima, 771-01 Japan
Tel: 0886-65-2888
Fax: 0886-37-1035.

FEATURES

SOURCE

Location/Qualifiers
1..392
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GEN-516H08"
/clone_lib="Human placenta polyA+ (TFujiwara)"
/tissue_type="placenta"

BASE COUNT 90 a 127 c 79 g 96 t

ORIGIN

Query Match 1.8%; Score 18; DB 157; Length 392;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 230 cccctgcacccccctg 247
|||||

Db 8 CCCCTGCATCACCCCTG 25

RESULT 39

AO253544/c

LOCUS

DEFINITION

HS-2227_A2_G04_MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2227 Col=8 Row=M, DNA sequence.

ACCESSION

AO253544

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 392)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

TITLE

JOURNAL
MEDLINE
COMMENT

Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 2227 row: M column: 8
Class: BAC ends

High quality sequence stop: 392.

Location/Qualifiers

FEATURES

SOURCE

1..392
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=2227 Col=8 Row=M"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC clones in E-Coli DH10B"
BASE COUNT 94 a 101 c 67 g 129 t 1 others

ORIGIN

Query Match 1.8%; Score 18; DB 226; Length 392;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 489 gccagctaagtgtgc 506
|||||

Db 154 GGCAGCTATATGTGTGC 137

RESULT 40

AV527255/c

LOCUS

DEFINITION

AV527255 Arabidopsis thaliana aboveground organs two to six-week old Arabidopsis thaliana cDNA clone AP23g91R 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 398)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

FEATURES

SOURCE

1..398
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="AP23g91R"
/clone_lib="Arabidopsis thaliana aboveground organs two to six-week old"
/tissue_type="aboveground organs"
/dev_stage="two to six-week old"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 95 a 102 c 94 g 107 t

ORIGIN

Query Match 1.8%; Score 18; DB 31; Length 398;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 930 caacatcagaaagggt 947
|||||

Db 35 CAACATCAGAAAGCGGT 18

RESULT 41

AV526564/c

LOCUS

DEFINITION

AV526564 Arabidopsis thaliana aboveground organs two to six-week old Arabidopsis thaliana cDNA clone AP216b09R 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 404)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and

JOURNAL MEDLINE
 20363093
 COMMENT The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamiizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/
 Location/Qualifiers
 1. 404
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone_11b="Arabidopsis thaliana aboveground organs two to six-week old"
 /tissue_type="aboveground organs"
 /dev_stage="two to six-week old"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"
 BASE COUNT 111 a 104 c 90 g 99 t
 ORIGIN

Query Match 1.8%; Score 18; DB 31; Length 404;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 930 caacatcacgaaggaggt 947
 ||||||||||||||||
 Db 141 CAACATCAGCAGGAGGT 124

RESULT 42
 LOCUS T21720/c 441 bp mRNA EST 07-JAN-1998
 DEFINITION 3728 Lambda-PR2 Arabidopsis thaliana cDNA clone 97C2177, mRNA sequence.
 ACCESSION T21720
 VERSION T21720.1 GI:2757230
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 1 (bases 1 to 441)
 Newman,T., debruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh ,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Kretzel ,E. and Somerville,C.
 Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 Plant Physiol. 106, 1241-1255 (1994)
 95148729
 On Jan 7, 1998 this sequence version replaced gi:914624.
 CONTACT: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313tcn@ldm.cl.msu.edu
 Seq primer: T7.
 Location/Qualifiers
 1. 441
 /organism="Arabidopsis thaliana"
 /strain="var Columbia"
 /db_xref="taxon:3702"
 /clone="97C2177"
 /clone_11b="Lambda-PR2"
 /note="Vector: lambda Zip-lox; Site_1: Sal; Site_2: Not; Lambda PRL2 is a cDNA library derived from equal

JOURNAL MEDLINE
 20363093
 COMMENT The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamiizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/
 Location/Qualifiers
 1. 404
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone_11b="Arabidopsis thaliana aboveground organs two to six-week old"
 /tissue_type="aboveground organs"
 /dev_stage="two to six-week old"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"
 BASE COUNT 111 a 104 c 90 g 99 t
 ORIGIN

Query Match 1.8%; Score 18; DB 188; Length 441;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 930 caacatcacgaaggaggt 947
 ||||||||||||||||
 Db 126 CAACATCAGCAGGAGGT 109

RESULT 43
 LOCUS H77242/c 443 bp mRNA EST 05-JAN-1998
 DEFINITION 17673 Lambda-PR2 Arabidopsis thaliana cDNA clone 205H2377, mRNA sequence.
 ACCESSION H77242
 VERSION H77242.1 GI:1054493
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 1 (bases 1 to 443)
 Newman,T., debruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh ,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Kretzel ,E. and Somerville,C.
 Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 Plant Physiol. 106, 1241-1255 (1994)
 95148729
 CONTACT: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313tcn@ldm.cl.msu.edu
 Seq primer: T7 dye primer.
 Location/Qualifiers
 1. 443
 /organism="Arabidopsis thaliana"
 /strain="var Columbia"
 /db_xref="taxon:3702"
 /clone="205H2377"
 /clone_11b="Lambda-PR2"
 /note="Vector: lambda Zip-lox; Site_1: Sal; Site_2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques. The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."
 BASE COUNT 122 a 104 c 95 g 105 t 17 others
 ORIGIN

Query Match 1.8%; Score 18; DB 158; Length 443;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 930 caacatcagaagggt 947
|||||
Db 153 CAACATCAGAGGGGT 136

RESULT 44
LOCUS AQ228766/c 446 bp DNA GSS 26-SEP-1998
DEFINITION HS.2020.A2.H08.T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2020 Col=16 Row=O, DNA sequence.
ACCESSION AQ228766
VERSION AQ228766.1 GI:3653995
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 446)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 2020 row: 0 column: 16
Class: BAC ends
High quality sequence stop: 446.
Location/Qualifiers
1.446
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=2020 Col=16 Row=O"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC clones in E-Coli DH10B"

BASE COUNT 72 a 131 c 103 g 138 t 2 others
ORIGIN

Query Match 1.8%; Score 18; DB 225; Length 446;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 816 agagctgcagcagcg 833
|||||
Db 232 AGAGCTGCAGCAGCG 215

RESULT 45
LOCUS N65546/c 461 bp mRNA EST 05-JAN-1998
DEFINITION 20586 Lambda-PRL2 Arabidopsis thaliana cDNA clone 231J27, mRNA sequence.
ACCESSION N65546
VERSION N65546.1 GI:1217172
KEYWORDS EST.
SOURCE Thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE 1 (bases 1 to 461)
AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Ohlrogge,J., Ralikel,N., Somerville,S., Thomashow,M., Rezel,E. and Somerville,C.
TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
JOURNAL Plant Physiol. 106, 1241-1255 (1994)
MEDLINE 95148729
COMMENT Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313c@lhm.cl.msu.edu
Seq primer: T7 dye primer.

FEATURES
source
1.461
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="var Columbia"
/db_xref="taxon:3702"
/clone="231J27"
/clone_lib="Lambda-PRL2"
/note="Vector: lambda Zip-Lox; Site.1: Sal; Site.2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques. The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."

BASE COUNT 118 a 107 c 100 g 119 t 17 others
ORIGIN

Query Match 1.8%; Score 18; DB 159; Length 461;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 930 caacatcagaagggt 947
|||||
Db 144 CAACATCAGAGGGGT 127

RESULT 46
LOCUS A1991951/c 463 bp mRNA EST 08-MAR-2000
DEFINITION ws43a04.x1 NCI-CGAP_Brn25 Homo sapiens cDNA IMAGE:249918 3' similar to contains element MSRI repetitive element ;, mRNA sequence.
ACCESSION A1991951
VERSION A1991951.1 GI:5838856
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 463)
AUTHORS NCI/NIHDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTCAP), Tumor Gene Index
UNPUBLISHED (1998)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/brp/image/image.html
Insert length: 565 Std Error: 0.00
Seq primer: -40UP from Gibco.

FEATURES

source

1. 463
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2499918"
/clone_1lb="NCI-CGAP_Brn25"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/note="Organ: brain; Vector: pRT73D-Pac (Pharmacia) with a
modified polylinker. Site.1: Not I; Site.2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TCTTACCAATCTGAAAGTGGAGCGCCGCGCATGATTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pRT73 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."

BASE COUNT 99 a 102 c 132 g 129 t 1 others

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred.No.1.1e+02; Length 463;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 230 ccccttcacacccctg 247

Db 151 ccccttcacacccctg 134

RESULT 47

BF612300

LOCUS

bf612300 463 bp mRNA EST 14-DEC-2000

daa17a11.y1 NICHD XGC lnl Xenopus laevis cDNA clone IMAGE:4056476

5' similar to SW:GSP_TORCA P56101 CYSTEINE STRING PROTEIN ;, mRNA

sequence.
BF612300

BF612300.1 GI:11782798

KEYWORDS

SOURCE

ORGANISM

African clawed frog.
Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;

Xenopodinae; Xenopus.

1 (bases 1 to 463)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: Xenopus clones from this library are available

through the I.M.A.G.E. Consortium/LLNL at: infoimage.llnl.gov

Seq primer: -40RP from Gibco

High quality sequence stop: 439.

FEATURES

source

1. 463
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="IMAGE:4056476"
/clone_1lb="NICHD XGC lnl"
/dev_stage="adult"

BASE COUNT

108 a 111 c 138 g 106 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred.No.1.1e+02; Length 463;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 453 cgggacacatgcgcgacc 470

Db 30 cgggacacatgcgcgacc 47

RESULT 48

BE811867

LOCUS

BE811867 465 bp mRNA EST 21-SEP-2000

PMO-AN0032-110700-001-a12 AN0032 Homo sapiens cDNA, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

sequence tags

/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pCMV-SpORF6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.6 kb. Constructed by Life
Technologies. Note: This is a Xenopus Gene Collection (XGC
) library."

FEATURES

source

1. 465
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="AN0032"
/dev_stage="Adult"

/note="Organ: amnion, normal; Vector: puc18; Site.1: SmaI;
Site.2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT

111 a 160 c 92 g 102 t

ORIGIN

Query Match

1.8%; Score 18; DB 140; Length 465;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 cccctgcatcccccctg 247
|||||

Db 138 CCCCTGCATCACCCCTG 155

RESULT 49

A2462258

LOCUS 474 bp DNA GSS 04-OCT-2000
DEFINITION M0269023F Mouse 10kb plasmid U06C1M library Mus musculus genomic
clone U06C1M0269023 F, DNA sequence.

ACCESSION

A2462258

VERSION

A2462258.1 GI:10620383

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. 474

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="U06C1M0269023"

/clone_lib="Mouse 10kb plasmid U06C1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PMD42ny; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adapter oligonucleotides were ligated to the blunt ends in high molar excess. The adapter DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g114732114|g114732114|g114732114) a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapterized mouse DNA was annealed to chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

ORIGIN

85 a 132 c 135 g 122 t

BASE COUNT

ORIGIN

85 a 132 c 135 g 122 t

BASE COUNT

ORIGIN

85 a 132 c 135 g 122 t

BASE COUNT

Query Match 1.8%; Score 18; DB 244; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 gaagcgcgcctcctg 282
|||||

Db 219 GCAGCCGCGCCTCCGCC 236

RESULT 50

A1466074

LOCUS 476 bp mRNA EST 09-MAR-1999
DEFINITION v039b02.y1 Soares_mammary-gland_NBMNG Mus musculus cDNA clone
IMAGE:1246155 5' similar to gb:U17574 Mouse mRNA for
proctosin-binding protein (MUSE);, mRNA sequence.

ACCESSION

A1466074

VERSION

A1466074.1 GI:4320104

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. 476

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:1246155"

/clone_lib="Soares_mammary-gland_NBMNG"

/sex="male"

/tissue_type="mammary gland"

/dev_stage="4 weeks"

/lab_host="DHI10B"

/note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5' TGTTACCAATCTGAAGTGGAGCGCGCGCAAGGTTTTTTTTTTTTTTTTTTT T 3'); double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

ORIGIN

104 a 132 c 134 g 106 t

BASE COUNT

ORIGIN

104 a 132 c 134 g 106 t

BASE COUNT

Query Match 1.8%; Score 18; DB 244; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 gaagcgcgcctcctg 282
|||||

Db 219 GCAGCCGCGCCTCCGCC 236

RESULT 50

A1466074

LOCUS 476 bp mRNA EST 09-MAR-1999
DEFINITION v039b02.y1 Soares_mammary-gland_NBMNG Mus musculus cDNA clone
IMAGE:1246155 5' similar to gb:U17574 Mouse mRNA for
proctosin-binding protein (MUSE);, mRNA sequence.

ACCESSION

A1466074

VERSION

A1466074.1 GI:4320104

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. 476

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:1246155"

/clone_lib="Soares_mammary-gland_NBMNG"

/sex="male"

/tissue_type="mammary gland"

/dev_stage="4 weeks"

/lab_host="DHI10B"

/note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5' TGTTACCAATCTGAAGTGGAGCGCGCGCAAGGTTTTTTTTTTTTTTTTTTT T 3'); double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

ORIGIN

104 a 132 c 134 g 106 t

BASE COUNT

ORIGIN

104 a 132 c 134 g 106 t

BASE COUNT

Mon Sep 24 09:36:19 2001

us-09-138-735-1_copy_1232_2207.oli5.rst

Page 30

Search completed: September 21, 2001, 03:21:19
Job time: 9395 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 16:24:12 : Search time 84.41 Seconds
(without alignments)
2188.931 Million cell updates/sec

Title: US-09-138-735-1_COPY_1232_2207

Perfect score: 976

Sequence: 1 cagtgacgagcgtaacgctt.....tgtgaacgagccattatgt 976

Scoring table:

OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 324599 seqs, 9465562 residues

Word size : 0

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

Issued_Patents_NA:*
1: /cgnl_7/ptodata/1/ina/5A.COMB.seq:*
2: /cgnl_7/ptodata/1/ina/5B.COMB.seq:*
3: /cgnl_7/ptodata/1/ina/6A.COMB.seq:*
4: /cgnl_7/ptodata/1/ina/6B.COMB.seq:*
5: /cgnl_7/ptodata/1/ina/PCRTUS.COMB.seq:*
6: /cgnl_7/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------------------|-------------------|
| 1 | 976 | 100.0 | 3402 | 1 US-08-480-917-1 | Sequence 1, Appl |
| 2 | 22 | 2.3 | 22 | 1 US-08-480-917-9 | Sequence 9, Appl |
| 3 | 21 | 2.2 | 21 | 1 US-08-480-917-10 | Sequence 10, Appl |
| 4 | 21 | 2.2 | 21 | 1 US-08-480-917-12 | Sequence 12, Appl |
| 5 | 19 | 1.9 | 11873 | 2 US-08-970-269A-32 | Sequence 32, Appl |
| 6 | 19 | 1.9 | 11878 | 2 US-08-970-269A-31 | Sequence 31, Appl |
| 7 | 19 | 1.9 | 11883 | 2 US-08-970-269A-28 | Sequence 28, Appl |
| 8 | 18 | 1.8 | 18 | 1 US-08-480-917-5 | Sequence 5, Appl |
| 9 | 18 | 1.8 | 18 | 1 US-08-480-917-8 | Sequence 8, Appl |
| 10 | 18 | 1.8 | 2520 | 2 US-08-454-557C-50 | Sequence 50, Appl |
| 11 | 18 | 1.8 | 2520 | 2 US-08-340-426D-50 | Sequence 50, Appl |
| 12 | 18 | 1.8 | 2520 | 2 US-08-450-673C-50 | Sequence 50, Appl |
| 13 | 18 | 1.8 | 2520 | 5 PCT-US95-17111A-50 | Sequence 50, Appl |
| 14 | 18 | 1.8 | 68750 | 3 US-09-335-409-1 | Sequence 1, Appl |
| 15 | 17 | 1.7 | 417 | 1 US-07-624-742-4 | Sequence 4, Appl |
| 16 | 17 | 1.7 | 673 | 2 US-08-692-787-4 | Sequence 4, Appl |
| 17 | 17 | 1.7 | 673 | 4 US-09-097-199-4 | Sequence 4, Appl |
| 18 | 17 | 1.7 | 2545 | 2 US-08-468-819-72 | Sequence 72, Appl |
| 19 | 17 | 1.7 | 3337 | 1 US-08-072-610-1 | Sequence 1, Appl |
| 20 | 17 | 1.7 | 3337 | 2 US-08-719-822B-1 | Sequence 1, Appl |
| 21 | 17 | 1.7 | 3337 | 4 US-09-092-458-1 | Sequence 1, Appl |
| 22 | 16 | 1.6 | 41 | 1 US-08-142-897-1 | Sequence 1, Appl |
| 23 | 16 | 1.6 | 737 | 4 US-08-998-416-1039 | Sequence 1039, Ap |
| 24 | 16 | 1.6 | 1185 | 2 US-08-628-039-9 | Sequence 9, Appl |
| 25 | 16 | 1.6 | 1185 | 3 US-08-912-205-9 | Sequence 9, Appl |
| 26 | 16 | 1.6 | 1185 | 4 US-09-440-400-9 | Sequence 9, Appl |
| 27 | 16 | 1.6 | 1314 | 2 US-09-031-059-2 | Sequence 2, Appl |

| | | | | | |
|-----|----|-----|-------|----------------------|--------------------|
| 28 | 16 | 1.6 | 1411 | 1 US-07-952-817-10 | Sequence 10, Appl |
| 29 | 16 | 1.6 | 1411 | 6 5210025-3 | Patent No. 5210025 |
| 30 | 16 | 1.6 | 1856 | 2 US-08-360-606B-29 | Sequence 29, Appl |
| 31 | 16 | 1.6 | 1864 | 2 US-08-673-388-9 | Sequence 9, Appl |
| 32 | 16 | 1.6 | 1864 | 2 US-08-614-877-9 | Sequence 9, Appl |
| 33 | 16 | 1.6 | 2190 | 3 US-08-508-761B-5 | Sequence 5, Appl |
| 34 | 16 | 1.6 | 2214 | 3 US-08-864-038A-1 | Sequence 1, Appl |
| 35 | 16 | 1.6 | 2310 | 3 US-09-036-987A-25 | Sequence 25, Appl |
| 36 | 16 | 1.6 | 3331 | 3 US-08-864-038A-2 | Sequence 2, Appl |
| 37 | 16 | 1.6 | 3331 | 3 US-08-864-038A-4 | Sequence 4, Appl |
| 38 | 16 | 1.6 | 4394 | 2 US-08-750-152A-1 | Sequence 1, Appl |
| 39 | 16 | 1.6 | 5117 | 5 PCT-US95-05512-1 | Sequence 1, Appl |
| 40 | 16 | 1.6 | 5117 | 5 PCT-US95-13749-3 | Sequence 3, Appl |
| 41 | 16 | 1.6 | 6545 | 5 US-08-056-200-93 | Sequence 93, Appl |
| 42 | 16 | 1.6 | 9551 | 2 US-08-800-644-93 | Sequence 93, Appl |
| 43 | 16 | 1.6 | 9551 | 2 US-09-009-913-1 | Sequence 1, Appl |
| 44 | 16 | 1.6 | 72928 | 3 US-08-633-779-2 | Sequence 2, Appl |
| 45 | 15 | 1.5 | 60 | 1 US-08-238-863-14 | Sequence 14, Appl |
| 46 | 15 | 1.5 | 81 | 1 US-08-443-407-14 | Sequence 14, Appl |
| 47 | 15 | 1.5 | 81 | 1 PCT-US95-05600-158 | Sequence 158, App |
| 48 | 15 | 1.5 | 81 | 5 US-08-994-946A-1 | Sequence 1, Appl |
| 49 | 15 | 1.5 | 325 | 3 US-09-383-586-2 | Sequence 2, Appl |
| 50 | 15 | 1.5 | 689 | 4 US-09-123-492A-3 | Sequence 3, Appl |
| 51 | 15 | 1.5 | 837 | 4 US-08-633-779-1 | Sequence 1, Appl |
| 52 | 15 | 1.5 | 838 | 1 US-09-154-874-1 | Sequence 1, Appl |
| 53 | 15 | 1.5 | 846 | 3 US-08-236-311-9 | Sequence 9, Appl |
| 54 | 15 | 1.5 | 1142 | 1 US-08-457-918-9 | Sequence 9, Appl |
| 55 | 15 | 1.5 | 1142 | 3 US-09-082-092-8 | Sequence 8, Appl |
| 56 | 15 | 1.5 | 1249 | 3 US-08-985-950-1 | Sequence 1, Appl |
| 57 | 15 | 1.5 | 1281 | 4 US-08-875-847B-1 | Sequence 1, Appl |
| 58 | 15 | 1.5 | 1313 | 1 US-08-176-427B-7 | Sequence 7, Appl |
| 59 | 15 | 1.5 | 1313 | 2 US-08-356-060A-4 | Sequence 4, Appl |
| 60 | 15 | 1.5 | 1313 | 4 US-08-855-910-5 | Sequence 5, Appl |
| 61 | 15 | 1.5 | 1438 | 3 US-08-651-136C-9 | Sequence 9, Appl |
| 62 | 15 | 1.5 | 1505 | 4 US-08-405-175A-10 | Sequence 10, Appl |
| 63 | 15 | 1.5 | 1566 | 2 US-08-211-718-8 | Sequence 8, Appl |
| 64 | 15 | 1.5 | 1608 | 2 US-07-959-941-1 | Sequence 1, Appl |
| 65 | 15 | 1.5 | 1632 | 1 US-08-259-924-1 | Sequence 1, Appl |
| 66 | 15 | 1.5 | 1632 | 1 US-08-699-103B-7 | Sequence 7, Appl |
| 67 | 15 | 1.5 | 1656 | 3 US-08-757-230A-1 | Sequence 1, Appl |
| 68 | 15 | 1.5 | 1715 | 4 PCT-US95-02315-1 | Sequence 1, Appl |
| 69 | 15 | 1.5 | 1715 | 5 US-09-348-443-1 | Sequence 1, Appl |
| 70 | 15 | 1.5 | 1763 | 4 US-08-173-508-5 | Sequence 5, Appl |
| 71 | 15 | 1.5 | 1777 | 2 US-08-265-310-5 | Sequence 5, Appl |
| 72 | 15 | 1.5 | 1777 | 3 US-08-951-742-5 | Sequence 5, Appl |
| 73 | 15 | 1.5 | 1777 | 3 US-08-934-494-1 | Sequence 1, Appl |
| 74 | 15 | 1.5 | 1869 | 3 US-09-143-068-1 | Sequence 1, Appl |
| 75 | 15 | 1.5 | 1869 | 3 US-09-350-268-1 | Sequence 1, Appl |
| 76 | 15 | 1.5 | 1869 | 4 US-08-642-406A-21 | Sequence 21, Appl |
| 77 | 15 | 1.5 | 1875 | 2 US-09-082-092-5 | Sequence 5, Appl |
| 78 | 15 | 1.5 | 1876 | 4 US-09-501-572-5 | Sequence 5, Appl |
| 79 | 15 | 1.5 | 1885 | 3 US-09-040-444-5 | Sequence 5, Appl |
| 80 | 15 | 1.5 | 1885 | 3 US-08-989-925-2 | Sequence 2, Appl |
| 81 | 15 | 1.5 | 1947 | 2 US-09-227-717-3 | Sequence 3, Appl |
| 82 | 15 | 1.5 | 2044 | 4 US-08-787-739-58 | Sequence 58, Appl |
| 83 | 15 | 1.5 | 2501 | 6 5457037-4 | Patent No. 5457037 |
| 84 | 15 | 1.5 | 2625 | 1 US-08-365-189-7 | Sequence 7, Appl |
| 85 | 15 | 1.5 | 2646 | 1 US-08-365-189-2 | Sequence 2, Appl |
| 86 | 15 | 1.5 | 2982 | 1 US-08-826-246-3 | Sequence 3, Appl |
| 87 | 15 | 1.5 | 2982 | 1 US-08-944-495-3 | Sequence 3, Appl |
| 88 | 15 | 1.5 | 3103 | 3 US-09-126-640-2 | Sequence 2, Appl |
| 89 | 15 | 1.5 | 3103 | 3 US-08-925-588-3 | Sequence 3, Appl |
| 90 | 15 | 1.5 | 3103 | 3 US-09-487-444-3 | Sequence 3, Appl |
| 91 | 15 | 1.5 | 3103 | 6 5457037-1 | Patent No. 5457037 |
| 92 | 15 | 1.5 | 3317 | 2 US-08-642-406A-20 | Sequence 20, Appl |
| 93 | 15 | 1.5 | 3317 | 2 US-08-434-000A-1 | Sequence 1, Appl |
| 94 | 15 | 1.5 | 3517 | 3 US-08-787-739-90 | Sequence 90, Appl |
| 95 | 15 | 1.5 | 3532 | 3 US-09-008-979A-7 | Sequence 7, Appl |
| 96 | 15 | 1.5 | 3532 | 3 US-09-460-618-7 | Sequence 7, Appl |
| 97 | 15 | 1.5 | 3981 | 4 US-09-298-367B-3 | Sequence 3, Appl |
| 98 | 15 | 1.5 | 4066 | | |
| 99 | 15 | 1.5 | | | |
| 100 | 15 | 1.5 | | | |

| | | | | | | | | | | | |
|-------|----|------|---|-------------------|--------------------|-------|----|------|---|-------------------|--------------------|
| 393 | 14 | 2335 | 2 | US-08-300-584-3 | Sequence 3, Appl1 | 466 | 14 | 3627 | 2 | US-09-025-583-7 | Sequence 7, Appl1 |
| 394 | 14 | 2336 | 4 | US-08-476-123-3 | Sequence 3, Appl1 | 467 | 14 | 3630 | 4 | US-09-378-255-5 | Sequence 5, Appl1 |
| 395 | 14 | 2335 | 4 | US-08-796-101-47 | Sequence 47, Appl1 | 468 | 14 | 3630 | 4 | US-09-141-212-1 | Sequence 1, Appl1 |
| C 396 | 14 | 2361 | 4 | US-09-011-197-1 | Sequence 1, Appl1 | 469 | 14 | 3630 | 4 | US-09-141-212-3 | Sequence 3, Appl1 |
| C 397 | 14 | 2409 | 3 | US-08-392-828C-1 | Sequence 1, Appl1 | 470 | 14 | 3630 | 4 | US-09-552-351-3 | Sequence 3, Appl1 |
| C 398 | 14 | 2409 | 3 | US-09-330-945-1 | Sequence 1, Appl1 | 471 | 14 | 3630 | 4 | US-09-251-372-5 | Sequence 5, Appl1 |
| 399 | 14 | 2420 | 2 | US-08-330-123A-3 | Sequence 3, Appl1 | C 472 | 14 | 3701 | 4 | US-08-845-258-10 | Sequence 10, Appl1 |
| 400 | 14 | 2425 | 2 | US-08-485-778-1 | Sequence 1, Appl1 | C 473 | 14 | 3701 | 4 | US-08-990-571-10 | Sequence 10, Appl1 |
| 401 | 14 | 2425 | 3 | US-08-520-550A-1 | Sequence 1, Appl1 | 474 | 14 | 3744 | 2 | US-08-348-955-16 | Sequence 16, Appl1 |
| 402 | 14 | 2426 | 1 | US-08-482-115B-3 | Sequence 3, Appl1 | 475 | 14 | 3744 | 2 | US-08-465-965-16 | Sequence 16, Appl1 |
| 403 | 14 | 2426 | 2 | US-08-660-678A-3 | Sequence 3, Appl1 | 476 | 14 | 3744 | 2 | US-08-465-966-16 | Sequence 16, Appl1 |
| 404 | 14 | 2426 | 2 | US-08-472-802C-4 | Sequence 4, Appl1 | 477 | 14 | 3607 | 1 | US-08-357-598-1 | Sequence 1, Appl1 |
| 405 | 14 | 2426 | 2 | US-08-714-482-1 | Sequence 1, Appl1 | 478 | 14 | 3607 | 2 | US-09-003-289-1 | Sequence 1, Appl1 |
| 406 | 14 | 2426 | 3 | US-08-998-443-3 | Sequence 3, Appl1 | 479 | 14 | 3607 | 5 | PCT-US95-16435-1 | Sequence 1, Appl1 |
| C 407 | 14 | 2455 | 3 | US-09-103-429A-1 | Sequence 3, Appl1 | C 480 | 14 | 4027 | 1 | US-08-348-143-4 | Sequence 3, Appl1 |
| C 408 | 14 | 2456 | 3 | US-08-813-150-5 | Sequence 5, Appl1 | C 481 | 14 | 4027 | 1 | US-08-348-143-4 | Sequence 3, Appl1 |
| C 409 | 14 | 2474 | 2 | US-08-666-392A-2 | Sequence 2, Appl1 | C 482 | 14 | 4027 | 1 | US-08-571-785-3 | Sequence 4, Appl1 |
| C 410 | 14 | 2529 | 3 | US-09-051-969A-5 | Sequence 5, Appl1 | C 483 | 14 | 4027 | 1 | US-08-571-785-4 | Sequence 4, Appl1 |
| C 411 | 14 | 2540 | 1 | US-08-446-919A-1 | Sequence 1, Appl1 | 484 | 14 | 4094 | 2 | US-08-841-483-5 | Sequence 5, Appl1 |
| C 412 | 14 | 2550 | 6 | 5258287-23 | Patent No. 5258287 | 485 | 14 | 4094 | 2 | US-09-382-912-5 | Sequence 5, Appl1 |
| 413 | 14 | 2609 | 4 | US-09-141-212-7 | Sequence 7, Appl1 | 486 | 14 | 4190 | 4 | US-08-924-345-1 | Sequence 1, Appl1 |
| 414 | 14 | 2625 | 4 | US-09-378-255-1 | Sequence 1, Appl1 | 487 | 14 | 4190 | 4 | US-08-781-891-205 | Sequence 205, App |
| 415 | 14 | 2669 | 4 | US-09-141-212-9 | Sequence 9, Appl1 | 488 | 14 | 4800 | 3 | US-09-106-638-1 | Sequence 1, Appl1 |
| 416 | 14 | 2693 | 4 | US-09-141-212-5 | Sequence 5, Appl1 | C 489 | 14 | 4849 | 2 | US-08-540-804-13 | Sequence 13, Appl1 |
| 417 | 14 | 2709 | 4 | US-09-251-372-1 | Sequence 1, Appl1 | C 490 | 14 | 4849 | 2 | US-08-218-263-13 | Sequence 13, Appl1 |
| 418 | 14 | 2733 | 4 | US-08-310-271-1 | Sequence 1, Appl1 | C 491 | 14 | 4849 | 2 | US-08-521-872-13 | Sequence 13, Appl1 |
| 419 | 14 | 2733 | 4 | US-09-032-742-9 | Sequence 9, Appl1 | C 492 | 14 | 4849 | 4 | US-08-590-399-13 | Sequence 13, Appl1 |
| 420 | 14 | 2764 | 2 | US-08-465-971B-1 | Sequence 1, Appl1 | C 493 | 14 | 4896 | 3 | US-09-053-866-1 | Sequence 1, Appl1 |
| 421 | 14 | 2816 | 1 | US-08-785-241-1 | Sequence 1, Appl1 | C 494 | 14 | 4896 | 2 | US-09-047-026A-1 | Sequence 1, Appl1 |
| C 422 | 14 | 2821 | 4 | US-09-103-429A-2 | Sequence 2, Appl1 | C 495 | 14 | 4822 | 2 | US-08-330-227-5 | Sequence 5, Appl1 |
| C 423 | 14 | 2887 | 3 | US-09-183-253-1 | Sequence 1, Appl1 | C 496 | 14 | 4822 | 2 | PCT-US93-13663-5 | Sequence 5, Appl1 |
| C 424 | 14 | 2943 | 2 | US-08-788-892-1 | Sequence 1, Appl1 | 497 | 14 | 5000 | 3 | US-09-104-070-1 | Sequence 1, Appl1 |
| C 425 | 14 | 2982 | 1 | US-08-348-143-2 | Sequence 1, Appl1 | C 498 | 14 | 5135 | 3 | US-08-369-822C-9 | Sequence 9, Appl1 |
| C 426 | 14 | 2982 | 1 | US-08-571-785-2 | Sequence 2, Appl1 | C 499 | 14 | 5135 | 3 | US-08-582-776C-9 | Sequence 9, Appl1 |
| C 427 | 14 | 2984 | 5 | PCT-US93-00893-1 | Sequence 2, Appl1 | C 500 | 14 | 5135 | 3 | US-08-582-776C-9 | Sequence 9, Appl1 |
| C 428 | 14 | 2984 | 5 | PCT-US93-00893-2 | Sequence 2, Appl1 | C 501 | 14 | 5135 | 3 | US-08-434-831B-9 | Sequence 9, Appl1 |
| C 429 | 14 | 2999 | 2 | US-09-014-969-8 | Sequence 8, Appl1 | 502 | 14 | 5183 | 3 | US-09-039-555B-18 | Sequence 18, Appl1 |
| C 430 | 14 | 3006 | 4 | US-09-552-351-1 | Sequence 1, Appl1 | 503 | 14 | 5275 | 4 | US-08-796-101-49 | Sequence 49, Appl1 |
| C 431 | 14 | 3018 | 3 | US-08-942-572-1 | Sequence 1, Appl1 | C 504 | 14 | 5275 | 4 | US-08-796-101-49 | Sequence 49, Appl1 |
| 432 | 14 | 3046 | 4 | US-09-378-255-3 | Sequence 3, Appl1 | 505 | 14 | 5470 | 6 | US-08-441-139-12 | Sequence 12, Appl1 |
| 433 | 14 | 3046 | 4 | US-08-726-725-1 | Sequence 1, Appl1 | C 506 | 14 | 5529 | 3 | US-08-869-696-1 | Sequence 1, Appl1 |
| 434 | 14 | 3057 | 1 | US-08-551-459-3 | Sequence 3, Appl1 | C 507 | 14 | 5559 | 1 | US-08-287-443-3 | Sequence 3, Appl1 |
| 435 | 14 | 3065 | 1 | US-08-156-866-1 | Sequence 1, Appl1 | C 508 | 14 | 5559 | 1 | US-08-459-701-3 | Sequence 3, Appl1 |
| C 436 | 14 | 3133 | 1 | US-08-162-809-1 | Sequence 1, Appl1 | C 509 | 14 | 5559 | 1 | US-08-460-296-3 | Sequence 3, Appl1 |
| C 437 | 14 | 3331 | 4 | US-08-195-152-1 | Sequence 1, Appl1 | C 510 | 14 | 5559 | 1 | US-08-459-174-3 | Sequence 3, Appl1 |
| C 438 | 14 | 3240 | 4 | US-09-262-773-7 | Sequence 7, Appl1 | C 511 | 14 | 5575 | 5 | PCT-US93-06300A-3 | Sequence 3, Appl1 |
| 439 | 14 | 3243 | 2 | US-08-239-276-6 | Sequence 6, Appl1 | C 512 | 14 | 5775 | 5 | US-08-306-691B-15 | Sequence 15, Appl1 |
| 440 | 14 | 3243 | 2 | US-08-468-579B-6 | Sequence 6, Appl1 | C 513 | 14 | 5775 | 5 | PCT-US93-06281-29 | Sequence 29, Appl1 |
| C 441 | 14 | 3243 | 3 | US-08-468-577B-6 | Sequence 6, Appl1 | 514 | 14 | 5855 | 1 | US-08-592-218A-20 | Sequence 20, Appl1 |
| C 442 | 14 | 3244 | 4 | US-09-262-773-3 | Sequence 3, Appl1 | 515 | 14 | 5855 | 3 | US-09-149-976-20 | Sequence 20, Appl1 |
| 443 | 14 | 3255 | 2 | US-08-916-917-11 | Sequence 11, Appl1 | 516 | 14 | 6008 | 1 | US-07-789-915A-5 | Sequence 5, Appl1 |
| C 444 | 14 | 3265 | 4 | US-09-225-170-11 | Sequence 11, Appl1 | 517 | 14 | 6008 | 1 | US-08-005-002C-5 | Sequence 5, Appl1 |
| C 445 | 14 | 3265 | 4 | US-09-262-773-5 | Sequence 5, Appl1 | 518 | 14 | 6008 | 1 | US-08-487-203A-5 | Sequence 5, Appl1 |
| C 446 | 14 | 3268 | 4 | US-09-262-773-1 | Sequence 1, Appl1 | 519 | 14 | 6036 | 1 | US-08-466-390-3 | Sequence 3, Appl1 |
| C 447 | 14 | 3288 | 2 | US-08-620-694A-1 | Sequence 1, Appl1 | 520 | 14 | 6036 | 1 | US-08-470-950-3 | Sequence 3, Appl1 |
| 448 | 14 | 3288 | 3 | US-09-022-255-1 | Sequence 1, Appl1 | 521 | 14 | 6306 | 1 | US-08-467-781-3 | Sequence 3, Appl1 |
| 449 | 14 | 3288 | 3 | US-09-022-696-1 | Sequence 1, Appl1 | 522 | 14 | 6306 | 1 | US-08-195-487-3 | Sequence 3, Appl1 |
| 450 | 14 | 3288 | 3 | US-08-978-776-11 | Sequence 1, Appl1 | 523 | 14 | 6306 | 2 | US-08-483-924-3 | Sequence 3, Appl1 |
| 451 | 14 | 3288 | 3 | US-09-022-253-1 | Sequence 1, Appl1 | 524 | 14 | 6306 | 5 | PCT-US93-06160-3 | Sequence 3, Appl1 |
| 452 | 14 | 3288 | 3 | US-09-022-260-1 | Sequence 1, Appl1 | 525 | 14 | 6317 | 1 | US-08-920-812-21 | Sequence 21, Appl1 |
| 453 | 14 | 3288 | 4 | US-09-022-259-1 | Sequence 1, Appl1 | 526 | 14 | 6317 | 1 | US-08-920-827-21 | Sequence 21, Appl1 |
| 454 | 14 | 3288 | 4 | US-09-022-257-1 | Sequence 1, Appl1 | 527 | 14 | 6317 | 1 | US-08-921-177-21 | Sequence 21, Appl1 |
| 455 | 14 | 3300 | 2 | US-08-928-692-29 | Sequence 29, Appl1 | 528 | 14 | 6317 | 1 | US-08-362-577C-21 | Sequence 21, Appl1 |
| C 456 | 14 | 3426 | 1 | US-08-234-939-1 | Sequence 1, Appl1 | 529 | 14 | 6317 | 2 | US-08-920-828-21 | Sequence 21, Appl1 |
| C 457 | 14 | 3426 | 1 | US-08-558-865-1 | Sequence 1, Appl1 | 530 | 14 | 6443 | 6 | 5183745-5 | Patent No. 5183745 |
| C 458 | 14 | 3426 | 3 | US-08-654-025-6 | Sequence 6, Appl1 | 531 | 14 | 6732 | 2 | US-09-127-670-5 | Sequence 5, Appl1 |
| C 459 | 14 | 3466 | 1 | US-08-551-425-5 | Sequence 5, Appl1 | 532 | 14 | 7032 | 3 | US-08-149-097D-24 | Sequence 24, Appl1 |
| 460 | 14 | 3546 | 1 | US-08-162-809-9 | Sequence 9, Appl1 | 533 | 14 | 7032 | 3 | US-08-949-386-24 | Sequence 24, Appl1 |
| 461 | 14 | 3591 | 1 | US-08-162-809-13 | Sequence 13, Appl1 | 534 | 14 | 7032 | 3 | US-08-450-562-24 | Sequence 24, Appl1 |
| C 462 | 14 | 3622 | 5 | US-08-306-691B-13 | Sequence 13, Appl1 | 535 | 14 | 7089 | 3 | US-08-949-386-25 | Sequence 25, Appl1 |
| C 463 | 14 | 3622 | 5 | PCT-US93-06251-24 | Sequence 24, Appl1 | 536 | 14 | 7881 | 3 | US-08-450-562-25 | Sequence 25, Appl1 |
| C 464 | 14 | 3627 | 1 | US-08-104-072B-6 | Sequence 6, Appl1 | 537 | 14 | 7881 | 2 | US-08-751-189-1 | Sequence 1, Appl1 |
| 465 | 14 | 3627 | 1 | US-08-351-413-7 | Sequence 7, Appl1 | 538 | 14 | 7881 | 2 | US-09-060-836-1 | Sequence 1, Appl1 |


```
RESULT 2
US-08-480-917-9
; Sequence 9, Application US/08480917
; Patent No. 5820864
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, glaucia
; APPLICANT: LESENECHAL, Mylene
; APPLICANT: JOLIVET, Michel
; TITLE OF INVENTION: NEW TRYPAOSOMA CRUZI ANTIGEN, AND GENE
; TITLE OF INVENTION: ENCODING THE LATTER; THEIR APPLICATION TO THE DETECTION OF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,917
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-480-917-9

Query Match 2.38; Score 22; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 35 cagccgacgtagctgctct 56
Db 1 CAGCCGACGCTAGCTGCTCT 22

RESULT 3
US-08-480-917-10/c
; Sequence 10, Application US/08480917
; Patent No. 5820864
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, glaucia
; APPLICANT: LESENECHAL, Mylene
; APPLICANT: JOLIVET, Michel
; TITLE OF INVENTION: NEW TRYPAOSOMA CRUZI ANTIGEN, AND GENE
; TITLE OF INVENTION: ENCODING THE LATTER; THEIR APPLICATION TO THE DETECTION OF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
```

```
ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,917
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-480-917-10

Query Match 2.28; Score 21; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 956 gtgcgaacgagccattatgt 976
Db 21 GTGCGAACGAGCCATTATGT 1

RESULT 4
US-08-480-917-12
; Sequence 12, Application US/08480917
; Patent No. 5820864
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, glaucia
; APPLICANT: LESENECHAL, Mylene
; APPLICANT: JOLIVET, Michel
; TITLE OF INVENTION: NEW TRYPAOSOMA CRUZI ANTIGEN, AND GENE
; TITLE OF INVENTION: ENCODING THE LATTER; THEIR APPLICATION TO THE DETECTION OF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,917
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
```


MEDIUM TYPE: Floppy disk
COMPUTER: Apple
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word for Macintosh
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/970,269A
FILING DATE: No. 5976803ember 14, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Adler Ph.D., Benjamin A.
REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5860
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-777-2321
TELEFAX: 713-777-6908
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 11883 bp
TYPE: nucleic acid
STRANDEDNESS: double stranded
TOPOLOGY: linear
MOLECULE TYPE: linear
DESCRIPTION: other nucleic acid
HYPOTHETICAL: no
ANTI-SENSE: no
ORIGINAL SOURCE:
US-08-970-269A-28

Query Match 1.9%; Score 19; DB 2; Length 11883;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 actctgcgcgtcgtcat 229
|||||
Db 6449 ACTTCTGCCGCTGCGCAT 6431

RESULT 8
US-08-480-917-5
Sequence 5, Application US/08480917
Patent No. 5820864
GENERAL INFORMATION:
APPLICANT: PARANHOS-BACCALA, Glaucia
APPLICANT: LESENECHAL, Mylene
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: NEW TRYPAOSOMA CRUZI ANTIGEN, AND GENE
TITLE OF INVENTION: ENCODING THE LATTER; THEIR APPLICATION TO THE DETECTION OF
TITLE OF INVENTION: CHAGAS DISEASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Oliff & Berridge
STREET: 700 South Washington Street, Suite 300
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,917
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 36400
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-480-917-5

Query Match 1.8%; Score 18; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 172 tcgggactgacggcgcg 189
|||||
Db 1 TCGGGACTGACGGCGCG 18

RESULT 9
US-08-480-917-8/C
Sequence 8, Application US/08480917
Patent No. 5820864
GENERAL INFORMATION:
APPLICANT: PARANHOS-BACCALA, Glaucia
APPLICANT: LESENECHAL, Mylene
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: NEW TRYPAOSOMA CRUZI ANTIGEN, AND GENE
TITLE OF INVENTION: ENCODING THE LATTER; THEIR APPLICATION TO THE DETECTION OF
TITLE OF INVENTION: CHAGAS DISEASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Oliff & Berridge
STREET: 700 South Washington Street, Suite 300
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,917
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 36400
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-480-917-8

Query Match 1.8%; Score 18; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 actctgcgcgtcgtcat 228
|||||
Db 18 ACTTCTGCCGCTGCGCAT 1

RESULT 10
US-08-454-557C-50
Sequence 50, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 2520 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-50

Query Match 1.8%; Score 18; DB 2; Length 2520;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 cccctgcatcaccctg 247
|||||
Db 298 CCCCTCATCACCCTG 315

RESULT 11
US-08-340-426D-50
Sequence 50, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 2520 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-50

Query Match 1.8%; Score 18; DB 2; Length 2520;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 cccctgcatcaccctg 247
|||||
Db 298 CCCCTCATCACCCTG 315

RESULT 12
US-08-450-673C-50
Sequence 50, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 2520 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-50

Query Match 1.8%; Score 18; DB 2; Length 2520;
Best Local Similarity 100.0%; Pred. No. 5.2;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 230 ccctgcatacccccctg 247
|||||
Db 298 CCCTGCATCACCCCTG 315

RESULT 13
PCT-US95-17111A-50
; Sequence 50, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Mands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/17111A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,426
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609,3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2540
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2520 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; PCT-US95-17111A-50

Query Match 1.8%; Score 18; DB 5; Length 2520;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 230 ccctgcatacccccctg 247
|||||
Db 298 CCCTGCATCACCCCTG 315

RESULT 14
US-09-335-409-1/c
; Sequence 1, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goelach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A

; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ. ID NOS: 30
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-335-409-1

Query Match 1.8%; Score 18; DB 3; Length 68750;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 517 tgcgtcccaagccgcg 534
|||||
Db 19020 TCGTCTCCACAGCGCG 19003

RESULT 15
US-07-624-742-4/c
; Sequence 4, Application US/07624742
; Patent No. 5236829
; GENERAL INFORMATION:
; APPLICANT: Farber, Joshua M.
; TITLE OF INVENTION: MONOKINE MIG INDUCED BY IFN-GAMMA
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie and Beckett
; STREET: 1001 G Street N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/624,742
; FILING DATE: 19901210
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107,29102
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 2025089100
; TELEFAX: 2025089299
; TELEX: 197430BBMB UT
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 417 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 40..414
US-07-624-742-4

Query Match 1.7%; Score 17; DB 1; Length 417;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 787 ctgaatcgtggttaga 803

DB 277 CTGAATCTGGTTAGA 261

RESULT 16
US-08-692-787-4

Sequence 4, Application US/08692787

Patent No. 5882864

GENERAL INFORMATION:

APPLICANT: An, Gang

APPLICANT: O'Hara, S. Mark

APPLICANT: Ralph, David

APPLICANT: Velti, Robert

TITLE OF INVENTION: BIOMARKERS AND TARGETS FOR DIAGNOSIS,

TITLE OF INVENTION: PROGNOSIS AND MANAGEMENT OF PROSTATE

NUMBER OF SEQUENCES: 82

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: United States of America

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/692,787

FILING DATE: Concurrently Herewith

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Corde, Timothy S.

REGISTRATION NUMBER: 38,414

REFERENCE/DOCKET NUMBER: UROC:012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (512) 474-7577

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 673 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-692-787-4

QY 537 gaaccaacagaggaag 553
DB 542 GAACCAACAGGAGAG 558

RESULT 17
US-09-097-199-4

Sequence 4, Application US/09097199

Patent No. 6218529

GENERAL INFORMATION:

APPLICANT: An, Gang

APPLICANT: O'Hara, S. Mark

APPLICANT: Ralph, David

APPLICANT: Velti, Robert

TITLE OF INVENTION: BIOMARKERS AND TARGETS FOR DIAGNOSIS,

TITLE OF INVENTION: PROGNOSIS AND MANAGEMENT OF PROSTATE DISEASE

NUMBER OF SEQUENCES: 87

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/097,199

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/692,787

FILING DATE: 31-JUL-1996

ATTORNEY/AGENT INFORMATION:

NAME: Nakashima, Richard A.

REGISTRATION NUMBER: P-42,023

REFERENCE/DOCKET NUMBER: UROC:018

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (512) 474-7577

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 673 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-097-199-4

Query Match 1.7%; Score 17; DB 4; Length 673;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaaccaacagaggaag 553
DB 542 GAACCAACAGGAGAG 558

RESULT 18

US-08-468-819-72/C

Sequence 72, Application US/08468819

Patent No. 5871723

GENERAL INFORMATION:

APPLICANT: Strieter, Robert M.

APPLICANT: Kunkel, Steven L.

TITLE OF INVENTION: CXc Chemokines as Regulators of

TITLE OF INVENTION: Angiogenesis

NUMBER OF SEQUENCES: 93

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: TX

COUNTRY: US

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/468,819

FILING DATE: Concurrently herewith

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Highlander, Steven L.

REGISTRATION NUMBER: 37,642

REFERENCE/DOCKET NUMBER: UMIC:003/HYL

TELECOMMUNICATION INFORMATION:

TELEPHONE: 512/418-3000
TELEFAX: 512/474-7477
TELEX: N/A
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 2545 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
US-08-468-819-72

Query Match 1.7%; Score 17; DB 2; Length 2545;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 787 ctgaatctgggttaga 803
|||||
DB 277 CTGAATCTGGGTTAGA 261

RESULT 19
US-08-072-610-1
Sequence 1, Application US/08072610
Patent No. 5532133
GENERAL INFORMATION:
APPLICANT: Barnwell, John
TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens,
TITLE OF INVENTION: Monoclonal Antibodies, and Diagnostic Assays
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby and Darby
STREET: 805 Third Ave.
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022-7513
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/072,610
FILING DATE: 19930602
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gogoris, Adda
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 5986/07686
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)527-7700
TELEFAX: (212)753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3337 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium vivax
IMMEDIATE SOURCE:
CLONE: PwMB3.3.1
US-08-072-610-1

Query Match

1.7%; Score 17; DB 1; Length 3337;

Best Local Similarity 100.0%; Pred. No. 17;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 aaccaacagagaaga 555
|||||
DB 2029 AACCAACAGAGAGAGAA 2045

RESULT 20
US-08-719-822B-1
Sequence 1, Application US/08719822B
Patent No. 5874527
GENERAL INFORMATION:
APPLICANT: Barnwell, John
TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby and Darby
STREET: 805 Third Ave.
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022-7513
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/719,822B
FILING DATE: 09/30/96
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gogoris, Adda
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 5986/17686US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)527-7700
TELEFAX: (212)753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3337 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium vivax
IMMEDIATE SOURCE:
CLONE: PwMB3.3.1
US-08-719-822B-1

Query Match 1.7%; Score 17; DB 2; Length 3337;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 aaccaacagagaaga 555
|||||
DB 2029 AACCAACAGAGAGAGAA 2045

RESULT 21
US-09-092-458-1
Sequence 1, Application US/09092458
Patent No. 6231861
GENERAL INFORMATION:
APPLICANT: Barnwell, John
TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens,
TITLE OF INVENTION: Monoclonal Antibodies, and Diagnostic Assays

PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 1039:
SEQUENCE CHARACTERISTICS:
LENGTH: 737 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: 1635RP
US-08-998-416-1039

Query Match 1.6%; Score 16; DB 4; Length 737;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 696 gtgcgtgagcgctg 711
|||||
DB 85 GTCGCTGAGCGCTG 100

RESULT 24
US-08-628-039-9/C
Sequence 9, Application US/08628039
Patent No. 5942660
GENERAL INFORMATION:
APPLICANT: Gruys, Kenneth J.
APPLICANT: Mitsky, Timothy A.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Slater, Steven C.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stark, David M.
APPLICANT: Hincbee, Maud A. W.
APPLICANT: Clemente, Thomas E.
APPLICANT: Connor-Ward, Dannelte V.
APPLICANT: Fedele, Mary J.
APPLICANT: Fry, Joyce E.
APPLICANT: Howe, Arlene R.
APPLICANT: Rozman, Renee J.
TITLE OF INVENTION: Methods of Optimizing Substrate Pools and
TITLE OF INVENTION: Biosynthesis of poly-B-hydroxybutyrate-co-poly-B-hydroxyvalerate
TITLE OF INVENTION: In Bacteria and Plants
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gary M. Bond, Monsanto Company, A35B
STREET: 800 No. 5942660th Lindbergh Boulevard
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63167
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/628,039
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE: 13-MAR-1996
ATTORNEY/AGENT INFORMATION:

NAME: Bond, Gary
REGISTRATION NUMBER: 29,283
REFERENCE/DOCKET NUMBER: 38-21(13585)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)694-3412
TELEFAX: (314)695-5435
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1185 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-628-039-9

Query Match 1.6%; Score 16; DB 2; Length 1185;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 506 ccagcgcgctgctg 521
|||||
DB 544 CCAGCGCGGCGCTG 529

RESULT 25
US-08-912-205-9/C
Sequence 9, Application US/08912205
Patent No. 6091002
GENERAL INFORMATION:
APPLICANT: Asrar, Jawed
APPLICANT: Mitsky, Timothy A.
APPLICANT: Shah, Devang T.
TITLE OF INVENTION: Polyhydroxyalkanoates of narrow molecular
TITLE OF INVENTION: Weight distribution prepared in transgenic plants
NUMBER OF SEQUENCES: 11
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/912,205
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/673,388
FILING DATE: 28-JUN-1996
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1185 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-912-205-9

Query Match 1.6%; Score 16; DB 3; Length 1185;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 506 ccagcgcgctgctg 521
|||||
DB 544 CCAGCGCGGCGCTG 529

RESULT 26
US-09-440-400-9/C
Sequence 9, Application US/09440400
Patent No. 6228623
GENERAL INFORMATION:
APPLICANT: Asrar, Jawed

APPLICANT: Mitsky, Timothy A.
APPLICANT: Shah, Devang T.
TITLE OF INVENTION: Polyhydroxyalkanoates of narrow molecular
TITLE OF INVENTION: weight distribution prepared in transgenic plants
NUMBER OF SEQUENCES: 11
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/440,400
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/912,205
FILING DATE:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1185 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-440-400-9

Query Match 1.6%; Score 16; DB 4; Length 1185;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 506 ccagcgagcctcgtc 521
|||||
DB 544 CCAGCGGCGCTCGTC 529

RESULT 27
US-09-031-059-2
Sequence 2, Application US/09031059
Patent No. 5948659
GENERAL INFORMATION:
APPLICANT: KATO, NOBUO
APPLICANT: SAKAI, YASUYOSHI
APPLICANT: TANI, YOSHIKI
APPLICANT: FUKUYA, HIROSHI
TITLE OF INVENTION: RECOMBINANT FRUCTOSYL AMINO ACID
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,059
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/899,172
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 0020-4253P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1314 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1311
US-09-031-059-2

Query Match 1.6%; Score 16; DB 2; Length 1314;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 351 taatgtaccacaag 366
|||||
DB 1281 TAATGTACCACCAAG 1296

RESULT 28
US-07-952-817-10
Sequence 10, Application US/07952817
Patent No. 5356796

GENERAL INFORMATION:
APPLICANT: Keller, John W.
TITLE OF INVENTION: A Repressor Protein and Gene for Regulating
TITLE OF INVENTION: Expression of Polypeptides and Its Use in the Preparation o
TITLE OF INVENTION: 2,2-Dialkylglycine Decarboxylase of Pseudomonas Cepacia
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finegan, Henderson, Farabow, Garrett &
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: US
ZIP: 20005-3315

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/952,817
FILING DATE: 19920928
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 01120,0002-01000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4400
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1411 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-952-817-10

Query Match 1.6%; Score 16; DB 1; Length 1411;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 459 caatggcgcactgtg 474
|||||
DB 372 CAATGGCGCCTGTC 387

RESULT 29
5210025-3
; Patent No. 5210025
; APPLICANT: KELLER, JOHN W.
; TITLE OF INVENTION: REPRESSOR PROTEIN GENE FOR REGULATING
; EXPRESSION OF POLYPEPTIDES AND ITS USE IN THE PREPARATION OF
; 2,2-DIALKYLGLYCINE DECARBOXYLASE OF PSEUDOMONAS CEPACIA
; NUMBER OF SEQUENCES: 18
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/501,814
; FILING DATE: 30-MAR-1990
; SEQ ID NO:3:
; LENGTH: 1411
5210025-3

Query Match 1.6%; Score 16; DB 6; Length 1411;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 459 caatgcccgcacctgtg 474
|||||
DB 372 caatgcccgcacctgtg 387

RESULT 30
US-08-360-606B-29/c
; Sequence 29, Application US/08360606B
; Patent No. 5919617
; GENERAL INFORMATION:
; APPLICANT: Jnanendra K. Bhattacharjee
; APPLICANT: Richard C. Garrod
; APPLICANT: Paul L. Skatrud
; APPLICANT: Robert P. Peery
; TITLE OF INVENTION: Methods and Reagents for
; TITLE OF INVENTION: Detecting Fungal Pathogens in a
; TITLE OF INVENTION: Biological Sample
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive Suite 3200
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,606B
; FILING DATE: December 21, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berghoff, Paul H.
; REGISTRATION NUMBER: 30,243
; REFERENCE/DOCKET NUMBER: 94,319
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)913-0001
; TELEFAX: (312)913-0002
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1856 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; HYPOHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Saccharomyces cerevisiae
US-08-360-606B-29

Query Match 1.6%; Score 16; DB 2; Length 1856;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 gtaatggcgcgcgcga 154
|||||
DB 167 gtaatggcgcgcgcga 152

RESULT 31
US-08-673-388-9/c
; Sequence 9, Application US/08673388
; Patent No. 5958745
; GENERAL INFORMATION:
; APPLICANT: Gruys, Kenneth J.
; APPLICANT: Mitsky, Timothy A.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stark, David M.
; APPLICANT: Hincbee, Maud A. W.
; APPLICANT: Clemente, Thomas E.
; APPLICANT: Connor-Ward, Dannelte V.
; APPLICANT: Fedele, Mary J.
; APPLICANT: Fry, Joyce E.
; APPLICANT: Howe, Arlene R.
; APPLICANT: Rozman, Renee J.
; TITLE OF INVENTION: Methods of Optimizing Substrate Pools and
; TITLE OF INVENTION: Biosynthesis of Poly-B-hydroxybutyrate-co-poly-B-hydroxyval
; TITLE OF INVENTION: in Bacteria and Plants
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Charles E. Cohen, Monsanto Company, BR4F
; STREET: 700 Chesterfield Parkway No. 5958745th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/673,388
; FILING DATE: 13-MAR-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Cohen, Charles E.
; REGISTRATION NUMBER: 34,565
; REFERENCE/DOCKET NUMBER: 38-21(13618)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6224
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1864 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-673-388-9

Query Match 1.6%; Score 16; DB 2; Length 1864;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 506 ccagcgcgcgcctgctc 521
|||||
DB 1073 CCAGCGCGGCCTGCTC 1058

RESULT 32
US-08-614-877-9/c
Sequence 9, Application US/08614877
Patent No. 5959179
GENERAL INFORMATION:
APPLICANT: Gruys, Kenneth J.
APPLICANT: Mitsky, Timothy A.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Slater, Steven C.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stark, David M.
APPLICANT: Hinchey, Maud A. W.
APPLICANT: Clemente, Thomas E.
APPLICANT: Connor-Ward, Dannelle V.
APPLICANT: Fedele, Mary J.
APPLICANT: Fry, Joyce E.
APPLICANT: Howe, Arlene R.
APPLICANT: Rozman, Renee J.
TITLE OF INVENTION: Methods of Optimizing Substrate Pools and
TITLE OF INVENTION: Biosynthesis of Poly-B-hydroxybutyrate-co-poly-B-hydroxyvalerate
TITLE OF INVENTION: In Bacteria and Plants
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Charles E. Cohen, Monsanto Company, B44F
STREET: 700 Chesterfield Parkway No. 5959179th
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/614,877
FILING DATE: 13-MAR-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Cohen, Charles E.
REGISTRATION NUMBER: 34,565
REFERENCE/DOCKET NUMBER: 38-21(10695)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6224
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1864 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-614-877-9

Query Match 1.6%; Score 16; DB 2; Length 1864;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 506 ccagcgcgagctcgcgc 521
|||||
DB 1073 CCAGCGCGCGCTCTC 1058

RESULT 33
US-08-508-761B-5
Sequence 5, Application US/08508761B
Patent No. 6027920
GENERAL INFORMATION:
APPLICANT: Joliff, Gwennael
APPLICANT: Guyonvarch, Arnel

APPLICANT: Purification, Relano
APPLICANT: Duchiron, Francis
APPLICANT: Renaud, Michel
TITLE OF INVENTION: System for Protein Expression and
TITLE OF INVENTION: Secretion Especially in Corynebacteria
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
STREET: 400 Seventh St. N.W.
CITY: Washington D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/508,761B
FILING DATE: 31-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91/09652
FILING DATE: 29-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91/09870
FILING DATE: 02-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Player, William E.
REGISTRATION NUMBER: 31,409
REFERENCE/DOCKET NUMBER: P58525NA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638-6666
TELEFAX: (202) 393-5350
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2190 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Corynebacterium Melassecola
FEATURE:
NAME/KEY: CDS
LOCATION: 573..1913
US-08-508-761B-5

Query Match 1.6%; Score 16; DB 3; Length 2190;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 864 aagctcgcggaatca 879
|||||
DB 1407 AAGCTCCGCGGAATCA 1422

RESULT 34
US-08-864-038A-1
Sequence 1, Application US/08864038A
Patent No. 6001592
GENERAL INFORMATION:
APPLICANT: Kunio NAKASHIMA et al.
TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR
TITLE OF INVENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBOD
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: 812-5 Hirano

STREET: Isshinden
CITY: Tsu-city
STATE: Mie-prefecture
COUNTRY: JAPAN
ZIP: 514-01
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: Microsoft Windows 95
SOFTWARE: Word Perfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/864,038A
FILING DATE: May 28, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-184459
FILING DATE: 15-July-1996
ATTORNEY/AGENT INFORMATION:
NAME: C. Bruce Hamburg
REGISTRATION NUMBER: 22,389
REFERENCE/DOCKET NUMBER: F-5610
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)986-2340
TELEFAX: (212)953-7733
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2214
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: Pinctada fucata
CELL TYPE: mantle epithelial cell
US-08-864-038A-1

Query Match 1.6%; Score 16; DB 3; Length 2214;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 215 ctgcgcgtctgcgc 230
|||||

Db 1544 CTGCGCGTCTGCATC 1559

RESULT 35
US-09-036-987A-25
Sequence 25, Application US/09036987A
Patent No. 6143526
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Turner, Jan R.
APPLICANT: Turney, Paul J.
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
TITLE OF INVENTION: Production
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dow Agrosciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/036,987A
FILING-DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R.
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 2310 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 88..1077
FEATURE:
NAME/KEY: CDS
LOCATION: 1165..1992
US-09-036-987A-25

Query Match 1.6%; Score 16; DB 3; Length 2310;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 395 cgacgaggtctacgac 410
|||||

Db 468 CGACGAGGTCTACGAC 483

RESULT 36
US-08-864-038A-2
Sequence 2, Application US/08864038A
Patent No. 6001592
GENERAL INFORMATION:
APPLICANT: Kunio NAKASHIMA et al.
TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE cDNA, VECTOR
TITLE OF INVENTION: CONTAINING SAID cDNA, HOST CELLS TRANSFORMED WITH SAID
TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBOD
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: 812-5 Hirano
STREET: Isshinden
CITY: Tsu-city
STATE: Mie-prefecture
COUNTRY: JAPAN
ZIP: 514-01
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: Microsoft Windows 95
SOFTWARE: Word Perfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/864,038A
FILING DATE: May 28, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-184459
FILING-DATE: 15-July-1996
ATTORNEY/AGENT INFORMATION:
NAME: C. Bruce Hamburg
REGISTRATION NUMBER: 22,389
REFERENCE/DOCKET NUMBER: F-5610
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)986-2340
TELEFAX: (212)953-7733
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 3331
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Pinctada fucata
CELL TYPE: mantle epithelial cell
FEATURE: mRNA
LOCATION: from 1 to 3331
IDENTIFICATION METHOD: E (by experiment)
US-08-864-038A-2

Query Match 1.6%; Score 16; DB 3; Length 3331;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 ctgccgcgtgcgtc 230
|||||
DB 1593 CTGCCGCTGTCATC 1608

RESULT 37
US-08-864-038A-4
Sequence 4, Application US/08864038A
Patent No. 6001592
GENERAL INFORMATION:
APPLICANT: Kunio NAKASHIMA et al.
TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE cDNA, VECTOR
TITLE OF INVENTION: CONTAINING SAID cDNA, HOST CELLS TRANSFORMED WITH SAID
TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: 812-5 Hirano
STREET: Isshinden
CITY: Tsu-city
STATE: Mie-prefecture
COUNTRY: JAPAN
ZIP: 514-01
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: Microsoft Windows 95
SOFTWARE: Word Perfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/864,038A
FILING DATE: May 28, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-184459
FILING DATE: 15-July-1996
ATTORNEY/AGENT INFORMATION:
NAME: C. Bruce Hamburg
REGISTRATION NUMBER: 22,389
REFERENCE/DOCKET NUMBER: F-5610
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)986-2340
TELEFAX: (212)953-7733
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3331
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Pinctada fucata
CELL TYPE: mantle epithelial cell
FEATURE:
NAME/KEY: CDS

LOCATION: from 50 to 2263
IDENTIFICATION METHOD: P (by similarity to some other pattern)
US-08-864-038A-4

Query Match 1.6%; Score 16; DB 3; Length 3331;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 ctgccgcgtgcgtc 230
|||||
DB 1593 CTGCCGCTGTCATC 1608

RESULT 38
US-08-750-152A-1/C
Sequence 1, Application US/08750152A
Patent No. 5977331
GENERAL INFORMATION:
APPLICANT: ASAKURA, YOKO
APPLICANT: KIMURA, EIICHIRO
APPLICANT: ABE, CHIZU
APPLICANT: KAWAHARA, YOSHIO
APPLICANT: NAKAMATSU, TSUYOSHI
TITLE OF INVENTION: ALPHA-KETOGLUTARATE DEHYDROGENASE GENE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22152
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/750,152A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4394 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Brevibacterium lactofermentum
STRAIN: ATCC13869
FEATURE:
NAME/KEY: CDS
LOCATION: 443..4213
FEATURE:
NAME/KEY: -35 signal
LOCATION: 281..287
FEATURE:
NAME/KEY: -10 signal
LOCATION: 307..312
FEATURE:

NAME/KEY: RBS
LOCATION: 421..428
FEATURE:
NAME/KEY: terminator
LOCATION: 4243..4281
US-08-750-152A-1

Query Match 1.6%; Score 16; DB 2; Length 4394;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 701 tggagcgcgtgagtc 716
|||||
Db 2624 TGGAGCGCGCTGAGTC 2609

RESULT 39
US-08-854-585-1/c
Sequence 1, Application US/08854585
Patent No. 6114140
GENERAL INFORMATION:
APPLICANT: Tonks, Nicholas K. and stman, Arne
TITLE OF INVENTION: Density Enhanced Protein Tyrosine Phosphatase
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, Suite 6300
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,585
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/237,940
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25,447
REFERENCE/DOCKET NUMBER: 27866/31954
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5117 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 350..4364
US-08-854-585-1

Query Match 1.6%; Score 16; DB 3; Length 5117;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 263 aggcagcgcgcctcc 278
|||||
Db 382 AGGCAGCGCGCCTCC 367

RESULT 40

PCT-US95-05512-1/c
Sequence 1, Application PC/TUS9505512
GENERAL INFORMATION:
APPLICANT: Tonks, Nicholas K. and stman, Arne
TITLE OF INVENTION: Density Enhanced Protein Tyrosine
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Borun
STREET: 233 South Wacker Drive, Suite 6300
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05512
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25,447
REFERENCE/DOCKET NUMBER: 27866/31954
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5117 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 350..4364
PCT-US95-05512-1

Query Match 1.6%; Score 16; DB 5; Length 5117;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 263 aggcagcgcgcctcc 278
|||||
Db 382 AGGCAGCGCGCCTCC 367

RESULT 41
PCT-US95-13749-3
Sequence 3, Application PC/TUS9513749
GENERAL INFORMATION:
APPLICANT: Amgen Inc.
TITLE OF INVENTION: COMPOSITIONS FOR INCREASED
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: PCT/US95/13749
FILING DATE:
CLASSIFICATION:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6545 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
MOLECULE TYPE: CDNA
TOPOLOGY: unknown
FEATURE:
NAME/KEY: CDS
LOCATION: 3630..4820
PCT-US95-13749-3

Query Match          1.6%; Score 16; DB 5; Length 6545;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 ggcgcgcgcgaacgcg 160
      |||
Db 5380 GCGACGCGCAACCG 5395

RESULT 42
US-08-056-200-93
Sequence 93, Application US/08056200
Patent No. 5616500
GENERAL INFORMATION:
APPLICANT: Steinert, Peter M.
APPLICANT: Lee, Seung-Chul
APPLICANT: Kim, In-Gyu
APPLICANT: Chung, Soo-Il
APPLICANT: Park, Sang-Chul
TITLE OF INVENTION: Trichoyalin and Transglutaminase-3 and
TITLE OF INVENTION: Methods of Using Same
NUMBER OF SEQUENCES: 117
CORRESPONDENCE ADDRESS:
ADDRESSER: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/056,200
FILING DATE: 30-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Pedrick, Michael F.
REGISTRATION NUMBER: 36,799
REFERENCE/DOCKET NUMBER: NIH054.001A
TELEPHONE: (714) 760-0404
TELEFAX: (714) 760-9502
INFORMATION FOR SEQ ID NO: 93:
SEQUENCE CHARACTERISTICS:
LENGTH: 9551 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: CDNA
TOPOLOGY: linear
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1507..1644

```

```

FEATURE:
NAME/KEY: Inton
LOCATION: 1645..2511
FEATURE:
NAME/KEY: CDS
LOCATION: 2512..8070
US-08-056-200-93

Query Match          1.6%; Score 16; DB 1; Length 9551;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 agggaaagcagcgcg 121
      |||
Db 6658 AGGAGAGAACGACGACG 6673

RESULT 43
US-08-800-644-93
Sequence 93, Application US/08800644
Patent No. 5958752
GENERAL INFORMATION:
APPLICANT: Steinert, Peter M.
APPLICANT: Lee, Seung-Chul
APPLICANT: Kim, In-Gyu
APPLICANT: Chung, Soo-Il
APPLICANT: Park, Sang-Chul
TITLE OF INVENTION: Trichoyalin and Transglutaminase-3 and
TITLE OF INVENTION: Methods of Using Same
NUMBER OF SEQUENCES: 117
CORRESPONDENCE ADDRESS:
ADDRESSER: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/800,644
FILING DATE: 14-FEB-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/056,200
FILING DATE: 30-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Pedrick, Michael F.
REGISTRATION NUMBER: 36,799
REFERENCE/DOCKET NUMBER: NIH054.001A
TELEPHONE: (714) 760-0404
TELEFAX: (714) 760-9502
INFORMATION FOR SEQ ID NO: 93:
SEQUENCE CHARACTERISTICS:
LENGTH: 9551 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: CDNA
TOPOLOGY: linear
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1507..1644
FEATURE:
NAME/KEY: Inton
LOCATION: 1645..2511

```

NAME/KEY: CDS
LOCATION: 2512..8070
US-08-800-644-93

Query Match 1.6%; Score 16; DB 2; Length 9551;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 106 aggaagagacagcagc 121
|||||
Db 6658 AGGAGAGACAGCAGC 6673

RESULT 44

US-09-009-913-1/C
Sequence 1, Application US/09009913
Patent No. 6087485
GENERAL INFORMATION:
APPLICANT: Axy's Pharmaceuticals, Inc.
TITLE OF INVENTION: Asthma Related Genes
NUMBER OF SEQUENCES: 339
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Ave, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,913
FILING DATE: 21-JAN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: SEQ-4P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3231
TELEFAX: 650-327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 72928 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-09-009-913-1

Query Match 1.6%; Score 16; DB 3; Length 72928;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 809 agatgacagactgca 824
|||||
Db 2985 AGATGACAGACTGCA 2970

RESULT 45

US-08-633-779-2/C
Sequence 2, Application US/08633779
Patent No. 5723332
GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: TRANSLATIONAL ENHANCER DNA
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderhye
STREET: 1100 No. 5723332th Glebe Road, 8th Floor
CITY: Arlington
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/633,779
FILING DATE: 19-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Leonard C. Mitchard
REGISTRATION NUMBER: 29,009
REFERENCE/DOCKET NUMBER: 604-381
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-633-779-2

Query Match 1.5%; Score 15; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 425 ctaccgccccgagc 439
|||||
Db 29 CTACGCCCCGAGC 15

RESULT 46
US-08-238-863-14
Sequence 14, Application US/08238863
Patent No. 5503978
GENERAL INFORMATION:
APPLICANT: SCHNEIDER, D. J., GOLD, L., AND FEIGON, J.
TITLE OF INVENTION: HIGH-AFFINITY ssDNA LIGANDS OF HIV-1
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: Beaton & Swanson, P.C.
STREET: 4582 South Ulster Street Parkway, Suite
STREET: #403
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80237
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/238,863
FILING DATE: 6-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/964,624
; FILING DATE: 21-OCTOBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 850-9900
; TELEFAX: (303) 850-9401
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 81
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-238-863-14

Query Match 1.5%; Score 15; DB 1; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 cccctgttcagcgc 256
Db 30 CCCCTGTTTCAGCGC 44

RESULT 47
US-08-443-407-14
; Sequence 14, Application US/08443407
; Patent No. 5786462
; GENERAL INFORMATION:
; APPLICANT: SCHNEIDER, D. J., GOLD, L., AND FEIGON, J.
; TITLE OF INVENTION: HIGH-AFFINITY sSDNA
; TITLE OF INVENTION: LIGANDS OF HIV-1 REVERSE
; TITLE OF INVENTION: TRANSCRIPTASE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/443,407
; FILING DATE:
; CLASSIFICATION:
; APPLICATION NUMBER: 08/238,863
; FILING DATE: 6-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/964,624
; FILING DATE: 21-OCTOBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX17/C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
```

```

; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 81
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-443-407-14

Query Match 1.5%; Score 15; DB 1; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 cccctgttcagcgc 256
Db 30 CCCCTGTTTCAGCGC 44

PCT-US95-05600-158
; Sequence 158, Application PC/TUS9505600
; GENERAL INFORMATION:
; APPLICANT: GOLD, LARRY
; APPLICANT: NIEUWLANDT, DAN
; APPLICANT: WECKER, MATTHEW
; APPLICANT: SCHNEIDER, DANIEL J.
; APPLICANT: FEIGON, JULI
; APPLICANT: ALLEN, PATRICK
; APPLICANT: SULENGER, BRUCE A.
; APPLICANT: DOUDNA, JENNIFER, A.
; TITLE OF INVENTION: HIGH-AFFINITY LIGANDS OF
; TITLE OF INVENTION: INSULIN RECEPTOR ANTIBODIES, TACHYKININ SUBSTANCE
; TITLE OF INVENTION: P, HIV INTEGRASE AND HIV-1 REVERSE TRANSCRIPTASE
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MG
; MEDIUM TYPE: Storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05600
; FILING DATE:
; CLASSIFICATION:
; APPLICATION NUMBER: 08/238,863
; FILING DATE: 06-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/248,632
; FILING DATE: 24-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/303,362
; FILING DATE: 09-SEPTEMBER-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/361,795
; FILING DATE: 21-DECEMBER-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/117,991
; FILING DATE: 08-SEPTEMBER-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/931,473
; FILING DATE: 17-AUGUST-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/964,624
; FILING DATE: 21-OCTOBER-1992
```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX17/PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 158:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 81 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US95-05600-158

```

```

Query Match          1.5%; Score 15; DB 5; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 242 cccctgttcagcgc 256
Db 30 CCCCTGTTTCAGCGC 44

```

```

RESULT 49
US-08-994-946A-1
; Sequence 1, Application US/08994946A
; Patent No. 6046317
; GENERAL INFORMATION:
; APPLICANT: Koulou, Markku
; APPLICANT: Karvonen, Matti
; APPLICANT: Pesonen, Ullamari
; APPLICANT: Uusitupa, Matti
; TITLE OF INVENTION: A DNA Molecule Encoding a Mutant
; TITLE OF INVENTION: Prepro-Neuropeptide Y, a Mutant Signal Peptide, and Uses
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
; STREET: 555 13th Street NW, Suite 701-E
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/994,946A
; FILING DATE: 19-DEC-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 2328-110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 base pairs

```

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-994-946A-1

```

```

Query Match          1.5%; Score 15; DB 3; Length 325;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 274 cctcctgcgcgcgcgc 288
Db 155 CCTCCTGCCGCGCGCG 169

```

```

RESULT 50
US-09-383-586-2/C
; Sequence 2, Application US/09383586
; Patent No. 6242419
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Mathew
; APPLICANT: Abernethy, Nevlin
; APPLICANT: Ormrod, Rene
; APPLICANT: Kumble, Anand
; APPLICANT: Kurison, Greg
; TITLE OF INVENTION: Compounds isolated from stromal cells
; TITLE OF INVENTION: and methods for their use
; FILE REFERENCE: 11000.1037c1
; CURRENT APPLICATION NUMBER: US/09/383,586
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 689
; TYPE: DNA
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35) ... (556)
; US-09-383-586-2

```

```

Query Match          1.5%; Score 15; DB 4; Length 689;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 900 cctgttcatttcgc 914
Db 218 CTTGTCTCATTTGC 204

```

```

Search completed: September 21, 2001, 16:27:14
Job time: 9005 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 16:15:12 : Search time 2753.51 Seconds
(without alignments)
5237.739 Million cell updates/sec

Title: US-09-138-735-1_COPY_1232_2207
Perfect score: 976
Sequence: 1 caggtacaggtacggtc.....tggtacaggtacattatgt 976

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 14155048 seqs, 7388405095 residues

Word size : 0

Total number of hits satisfying chosen parameters: 28310096

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

Pending_Patents_NA_Main:*

1: /cgnl_7/ptodata/1/pna/PCrUS_COMB.seq:*
2: /cgnl_7/ptodata/1/pna/US06_COMB.seq:*
3: /cgnl_7/ptodata/1/pna/US07_COMB.seq:*
4: /cgnl_7/ptodata/1/pna/US08_COMB.seq:*
5: /cgnl_7/ptodata/1/pna/US081_COMB.seq:*
6: /cgnl_7/ptodata/1/pna/US082_COMB.seq:*
7: /cgnl_7/ptodata/1/pna/US083_COMB.seq:*
8: /cgnl_7/ptodata/1/pna/US084_COMB.seq:*
9: /cgnl_7/ptodata/1/pna/US085_COMB.seq:*
10: /cgnl_7/ptodata/1/pna/US086_COMB.seq:*
11: /cgnl_7/ptodata/1/pna/US087_COMB.seq:*
12: /cgnl_7/ptodata/1/pna/US088_COMB.seq:*
13: /cgnl_7/ptodata/1/pna/US089_COMB.seq:*
14: /cgnl_7/ptodata/1/pna/US090_COMB.seq:*
15: /cgnl_7/ptodata/1/pna/US091_COMB.seq:*
16: /cgnl_7/ptodata/1/pna/US092_COMB.seq:*
17: /cgnl_7/ptodata/1/pna/US093_COMB.seq:*
18: /cgnl_7/ptodata/1/pna/US094_COMB.seq:*
19: /cgnl_7/ptodata/1/pna/US095_COMB.seq:*
20: /cgnl_7/ptodata/1/pna/US096_COMB.seq:*
21: /cgnl_7/ptodata/1/pna/US097_COMB.seq:*
22: /cgnl_7/ptodata/1/pna/US098_COMB.seq:*
23: /cgnl_7/ptodata/1/pna/US099_COMB.seq:*
24: /cgnl_7/ptodata/1/pna/US100_COMB.seq:*
25: /cgnl_7/ptodata/1/pna/US101_COMB.seq:*
26: /cgnl_7/ptodata/1/pna/US102_COMB.seq:*
27: /cgnl_7/ptodata/1/pna/US103_COMB.seq:*
28: /cgnl_7/ptodata/1/pna/US104_COMB.seq:*
29: /cgnl_7/ptodata/1/pna/US105_COMB.seq:*
30: /cgnl_7/ptodata/1/pna/US106_COMB.seq:*
31: /cgnl_7/ptodata/1/pna/US107_COMB.seq:*
32: /cgnl_7/ptodata/1/pna/US108_COMB.seq:*
33: /cgnl_7/ptodata/1/pna/US109_COMB.seq:*
34: /cgnl_7/ptodata/1/pna/US110_COMB.seq:*
35: /cgnl_7/ptodata/1/pna/US111_COMB.seq:*
36: /cgnl_7/ptodata/1/pna/US112_COMB.seq:*
37: /cgnl_7/ptodata/1/pna/US113_COMB.seq:*
38: /cgnl_7/ptodata/1/pna/US114_COMB.seq:*
39: /cgnl_7/ptodata/1/pna/US115_COMB.seq:*
40: /cgnl_7/ptodata/1/pna/US116_COMB.seq:*
41: /cgnl_7/ptodata/1/pna/US117_COMB.seq:*
42: /cgnl_7/ptodata/1/pna/US118_COMB.seq:*
43: /cgnl_7/ptodata/1/pna/US119_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1 | 976 | 100.0 | 3402 | US-08-988-242-1 | Sequence 1, Appl1 |
| 2 | 976 | 100.0 | 3402 | US-09-138-736-1 | Sequence 1, Appl1 |
| 3 | 22 | 2.3 | 22 | US-08-988-242-5 | Sequence 5, Appl1 |
| 4 | 22 | 2.3 | 22 | US-09-138-736-9 | Sequence 9, Appl1 |
| 5 | 22 | 2.3 | 366 | US-09-606-977-3519 | Sequence 3519, Ap |
| 6 | 21 | 2.2 | 21 | US-08-988-242-6 | Sequence 6, Appl1 |
| 7 | 21 | 2.2 | 21 | US-08-988-242-7 | Sequence 7, Appl1 |
| 8 | 21 | 2.2 | 21 | US-09-138-736-10 | Sequence 10, Appl1 |
| 9 | 21 | 2.2 | 21 | US-09-138-736-11 | Sequence 11, Appl1 |
| 10 | 20 | 2.0 | 391 | US-09-332-782-18690 | Sequence 18690, A |
| 11 | 20 | 2.0 | 391 | US-09-515-694-18690 | Sequence 18690, A |
| 12 | 20 | 2.0 | 403 | US-09-332-782-12269 | Sequence 12269, A |
| 13 | 20 | 2.0 | 403 | US-09-515-694-12269 | Sequence 12269, A |
| 14 | 20 | 2.0 | 408 | US-09-332-782-12260 | Sequence 12260, A |
| 15 | 20 | 2.0 | 408 | US-09-515-694-12260 | Sequence 12260, A |
| 16 | 20 | 2.0 | 718 | US-09-654-617-126156 | Sequence 126156, A |
| 17 | 20 | 2.0 | 718 | US-09-684-016-126156 | Sequence 126156, A |
| 18 | 20 | 2.0 | 5679 | US-09-620-392-64793 | Sequence 64793, A |
| 19 | 20 | 2.0 | 5679 | US-09-702-134-21010 | Sequence 21010, A |
| 20 | 20 | 2.0 | 11335 | US-60-173-464-11335 | Sequence 11353, A |
| 21 | 20 | 2.0 | 11335 | US-60-191-637-13953 | Sequence 13953, A |
| 22 | 20 | 2.0 | 11335 | US-60-191-681-10986 | Sequence 10986, A |
| 23 | 20 | 2.0 | 11338 | US-60-167-217-13998 | Sequence 13998, A |
| 24 | 20 | 2.0 | 91470 | US-09-534-859-19 | Sequence 19, Appl1 |
| 25 | 20 | 2.0 | 226706 | US-09-528-237A-1285 | Sequence 1285, Ap |
| 26 | 19 | 1.9 | 718 | US-60-182-316-2055 | Sequence 2055, Ap |
| 27 | 19 | 1.9 | 9855 | US-09-620-392-32691 | Sequence 32691, A |
| 28 | 19 | 1.9 | 9895 | US-09-620-392-67713 | Sequence 67713, A |
| 29 | 19 | 1.9 | 11873 | US-09-407-562-32 | Sequence 32, Appl1 |
| 30 | 19 | 1.9 | 11878 | US-09-407-562-31 | Sequence 31, Appl1 |
| 31 | 19 | 1.9 | 11883 | US-09-420-362-28 | Sequence 28, Appl1 |
| 32 | 19 | 1.9 | 19505 | US-09-620-392-10632 | Sequence 10632, A |
| 33 | 19 | 1.9 | 25670 | US-09-702-134-5295 | Sequence 5295, Ap |
| 34 | 19 | 1.9 | 31556 | US-09-702-134-5301 | Sequence 5301, Ap |
| 35 | 18 | 1.8 | 159989 | US-08-988-242-4 | Sequence 4, Appl1 |
| 36 | 18 | 1.8 | 18 | US-09-138-736-5 | Sequence 5, Appl1 |
| 37 | 18 | 1.8 | 18 | US-09-366-691A-1504 | Sequence 1504, Ap |
| 38 | 18 | 1.8 | 70 | US-09-076-667-928 | Sequence 928, App |
| 39 | 18 | 1.8 | 85 | US-09-540-229-15688 | Sequence 15688, A |
| 40 | 18 | 1.8 | 85 | US-60-048-002-928 | Sequence 928, App |
| 41 | 18 | 1.8 | 180 | US-09-513-991-2039 | Sequence 2039, Ap |
| 42 | 18 | 1.8 | 194 | US-09-654-617-141143 | Sequence 141143, A |
| 43 | 18 | 1.8 | 194 | US-09-684-016-141143 | Sequence 141143, A |
| 44 | 18 | 1.8 | 196 | US-09-654-617-162528 | Sequence 162528, A |
| 45 | 18 | 1.8 | 25 | | |
| 46 | 18 | 1.8 | 25 | | |

Db 1412 GACGGGAGTACAGTATCAATCAATACGACTTCTGCCGCTGTCATCCCTCATCA 1471
QY 241 cccctgttcaagcgcccaagcgagcgccctctctccgagcgagcgatcgctag 300
Db 1472 CCCCTGTTCAGCCGCCAGCCAGGACCGCGCTCTCTGCGCGCGCGGATCGCTGAG 1531
QY 301 ccgacgctgggagcaagatcatctgaatctgaatctgaatctgaatctgaatctgaat 360
Db 1532 CCGCAGCTGGGAGAGCAATATCTTCTAATCTAATCTAATCTAATCTAATCTAATCTAAT 1591
QY 361 caaagagcgctcgtcagcactggagcgccgagcgagcgagcgagcgagcgagcgagcg 420
Db 1592 CAAAGAGCGCTGTCAGACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1651
QY 421 acgactaccgcccgcagcaagatccatcaagcgagcgagcgagcgagcgagcgagcgag 480
Db 1652 ACGACTACCGCGCCCGCAGCAAGTCCATACGGGCGCAATGGCGCGAGCTGTGACGGCT 1711
QY 481 ggaattgctgagcactaagtagtggtagcagcgcgccctctctccacagcgcgcgagaa 540
Db 1712 GGATTGGTGGAGCTAATAGTGTGTCAGCGCGCGCTCTGTCCTCCACAGCGCGCGCGGAAA 1771
QY 541 ccaacagagaagaagaagcgctcgcgagcgatgtgaacgagcgagcgagcgagcgagcg 600
Db 1772 CCAACAGAGAGAGAGAGAGAGCGCTCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1831
QY 601 aacagcgcgcgcttcaaacgctctcccgagcgagcgagcgagcgagcgagcgagcgag 660
Db 1832 ACACGCCGCGGCTTCAACAGCGCTCTCTCCGCGCGCGCGCGCGCGCGCGCGCGCTTTGGCG 1891
QY 661 gcaagcagataccagcagcagcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 720
Db 1892 GCAGCAGATATACAGCGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 1951
QY 721 ataacaacacgctcgtgagcttgaagtgctcctgaacacacacacacacacacacacac 780
Db 1952 ATAACAACACGCTCGCGGCTCTCTGAGTGTCTCCGCGCGCGCGCGCGCGCGCGCGCT 2011
QY 781 caactctgaatctgagtgtagagcagacatgacagagcgagcgagcgagcgagcgagcg 840
Db 2012 CACTTCTGATCTGGGTTTGAAGGACACAGATGACAGAGCTGACAGAGCGCGCTCAACA 2071
QY 841 ccgacacacagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 900
Db 2072 CCGCAACACAGCGCGAGAGACACAGCTCCGCGAGATCATCGCTTTTGAAGCGATACAC 2131
QY 901 ctgttctcatttgagatctcctctctcgcacacacacacacacacacacacacacacac 960
Db 2132 CTGTTCATTTGCGGATTCCTCTCTCTGCAACATCAGAGGCGGCTGAGCGTGTG 2191
QY 961 aagagcgcatatgt 976
Db 2192 AACGAGCGCATATATGT 2207

RESULT 2
US-09-138-736-1
Sequence 1, Application US/09138736
GENERAL INFORMATION:
APPLICANT: PARANHOS-BACALA, Glaucia
APPLICANT: LESENECHAL, Mylene
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: NEW TRYPAOSOMA CRUZI ANTIGEN, AND GENE
ENCODING THE LATTER: THEIR APPLICATION TO THE DETECTION OF
TITLE OF INVENTION: CHAGAS DISEASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESS: 700 South Washington Street, Suite 300
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/138,736
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/480,917
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Beridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 36400
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3402 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-138-736-1

Query Match: 100.0%; Score 976; DB 15; Length 3402;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 976; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 caaggtacagcgtaagcgcttcttctcaatcgttaacgcgaagcgtagctgctgct 60
Db 1232 CAGGTACACCGTAAAGCGCTTTGCTTCAATCTTACAGCGAGCGAGTGTGCTGTGCT 1291
QY 61 gacatgctcattcgaatgacatcttccatctcgcgctcctccgcgaggaagaacagcg 120
Db 1292 GACATGCTCATTCGATTGACGATCTTCCATCTCCGCTCTCCCGAGGAGAGACGACG 1351
QY 121 ccagggccaaaac 180
Db 1352 CCAGGCCAAAACATCGGTAGTGGGAGCGGAGAAACCGGGGTGTGCTCGGGGCACT 1411
QY 181 gacgcgagcgagtagcagtcatacacaacacacacacacacacacacacacacacacac 240
Db 1412 GACGGGCGAGTAGCAGTATCAATACGACTTGTGCGCTGTCATCCCTCATCA 1471
QY 241 cccctgttcaagcgcccaagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 300
Db 1472 CCCCTGTTCAGCGCCAGGCGAGCGCGCTCTCTGCGCGCGCGCGCGCGCGCGCTGAG 1531
QY 301 ccgacgctgggagcaagatcatctgaatctgaatctgaatctgaatctgaatctgaatct 360
Db 1532 CCGCAGCTGGGAGAGCAATATCTTCTAATCTAATCTAATCTAATCTAATCTAATCTAAT 1591
QY 361 caaagagcgctcgtcagcactggagcgccgagcgagcgagcgagcgagcgagcgagcg 420
Db 1592 CAAAGAGCGCTGTCAGACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1651
QY 421 acgactaccgcccgcagcaagatccatcaagcgagcgagcgagcgagcgagcgagcgag 480
Db 1652 ACGACTACCGCGCCCGCAGCAAGTCCATACGGGCGCAATGGCGCGAGCTGTGACGGCT 1711
QY 481 ggaattgctgagcactaagtagtggtagcagcgcgccctctctccacagcgcgcgagaa 540
Db 1712 GGATTGGTGGAGCTAATAGTGTGTCAGCGCGCGCTCTGTCCTCCACAGCGCGCGGAAA 1771
QY 541 ccaacagagaagaagaagcgctcgcgagcgatgtgaacgagcgagcgagcgagcgagcg 600
Db 1772 CCAACAGAGAGAGAGAGAGCGCTCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCTTTGGCG 1831

QY 601 acacgccccgagcttcaacaagcctctctccgcagcgcccaacgagtgagctttgagcg 660
| | | | |
Db 1832 ACAGCGCCCGCGCTTACACAGCGCTCTCTCCGACGCGCCCAACGAGTGGCGTTTGCGC 1891
QY 661 gacgaggtatacaagtcgagggcgaggttcacagtcgctgtgagcggtcggtccgtc 720
| | | | |
Db 1892 GCAGCAGTATACAGTCGGAGGCGAGGTCATCAGTCGTCGAGCGCGGTGAGTCCGTC 1951
QY 721 ataacacaacgctcgcggttcgaagttgctccctacacccattcgaagaacatgaa 780
| | | | |
Db 1952 ATACACACAGCGCTCGGCTTGTGAAGTTCCTCCCTACACCAATTCGAAACAGACATGAA 2011
QY 781 caactctcgaatctgggttttagagagcagatgacagagctgacagagcgctccaaca 840
| | | | |
Db 2012 CAACCTCTGAATCTGGTTTGAAGAGCACAGATGACAGAGCTGCAGACCGCTCCCAACA 2071
QY 841 cgcgaacacagcgagagagacacagctcgcgaatcaccgtgtttgagaagtaacc 900
| | | | |
Db 2072 CCGCAACACAGCGAGAGAGACACAGCTCCGCAAAATCATCCGTTGTGAGAGCTACACC 2131
QY 901 ctgtctcattgaggttcctcctctcgcgaacatcacagagggagtgagcgctgtgtg 960
| | | | |
Db 2132 CTGTCTCTCTGTCGGATTCCTCTCTCGACATCATCGAAGGGGTGAGAGCTGTGTG 2191
QY 961 aacgagggcattatgt 976
| | | | |
Db 2192 AACGAGCGCATATATGT 2207

RESULT 3
US-08-988-242-5
; Sequence 5, Application US/08988242
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: LESENECHAL, MYLENE
; APPLICANT: JOLIVET, MICHEL
; APPLICANT: MANDRAND, BERNARD
; TITLE OF INVENTION: NEW TRYPANOSOMA CRUZI ANTIGEN, GENE
; TITLE OF INVENTION: ENCODING THEREFOR, AND METHODS OF DETECTING AND TREATING
; TITLE OF INVENTION: CHAGAS DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIVE & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/988,242
; FILING DATE: 10-DEC-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-988-242-5

Query Match 2.3%; Score 22; DB 13; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 cagcgcaggtagctgcgtcct 56
| | | | |
Db 1 CAGCGCAGGTAGCTGCGTCT 22

RESULT 4
US-09-138-736-9
; Sequence 9, Application US/09138736
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: LESENECHAL, MYLENE
; APPLICANT: JOLIVET, MICHEL
; TITLE OF INVENTION: NEW TRYPANOSOMA CRUZI ANTIGEN, AND GENE
; TITLE OF INVENTION: ENCODING THE LATTER; THEIR APPLICATION TO THE DETECTION OF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Olive & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/138,736
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,917
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-138-736-9

Query Match 2.3%; Score 22; DB 15; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 cagcgcaggtagctgcgtcct 56
| | | | |
Db 1 CAGCGCAGGTAGCTGCGTCT 22

RESULT 5
US-09-606-977-3519/C
; Sequence 3519, Application US/09606977
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated W


```
FILE REFERENCE: 38-21(15877)B
; CURRENT APPLICATION NUMBER: US/09/606,977
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 82359
; SEQ ID NO 3519
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: unsure at all n locations
US-09-606-977-3519

Query Match 2.3%; Score 22; DB 23; Length 366;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 810 gatgacagagctgcagcagagc 831
|||||
Db 269 gatgacagagctgcagcagagc 248

RESULT 6
US-08-988-242-6/c
; Sequence 6, Application US/08988242
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: LESENECHAL, MYLENE
; APPLICANT: JOLIVET, MICHEL
; APPLICANT: MANDRAND, BERNARD
; TITLE OF INVENTION: NEW TRYPANOSOMA CRUZI ANTIGEN, GENE
; TITLE OF INVENTION: ENCODING THEREFOR, AND METHODS OF DETECTING AND TREATING
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/988,242
; FILING DATE: 10-DEC-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-988-242-6

Query Match 2.2%; Score 21; DB 13; Length 21;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 956 gtgtgaacgagcattatgt 976
|||||
Db 21 gtgtgaacgagcattatgt 1
```

```
RESULT 7
US-08-988-242-7
; Sequence 7, Application US/08988242
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: LESENECHAL, MYLENE
; APPLICANT: JOLIVET, MICHEL
; APPLICANT: MANDRAND, BERNARD
; TITLE OF INVENTION: NEW TRYPANOSOMA CRUZI ANTIGEN, GENE
; TITLE OF INVENTION: ENCODING THEREFOR, AND METHODS OF DETECTING AND TREATING
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/988,242
; FILING DATE: 10-DEC-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-988-242-7

Query Match 2.2%; Score 21; DB 13; Length 21;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 766 cgaagagacattgaacaatt 786
|||||
Db 1 cgaagagacattgaacaatt 21

RESULT 8
US-09-138-736-10/c
; Sequence 10, Application US/09138736
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: LESENECHAL, MYLENE
; APPLICANT: JOLIVET, MICHEL
; APPLICANT: MANDRAND, BERNARD
; TITLE OF INVENTION: NEW TRYPANOSOMA CRUZI ANTIGEN, AND GENE
; TITLE OF INVENTION: ENCODING THE LATTER, THEIR APPLICATION TO THE DETECTION OF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
```

```

; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/138,736
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,917
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-138-736-10

Query Match      2.2%; Score 21; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 956 gtgtgaacgagaccattgt 976
    |||||||||||||||||||
DB 21 ctctgaacgagccattatct 1

RESULT 9
US-09-138-736-12
; Sequence 12, Application US/09138736
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: LESENECHAL, Mylene
; APPLICANT: JOLIVET, Michel
; TITLE OF INVENTION: NEW TRYPAOSOMA CRUZI ANTIGEN, AND GENE
; TITLE OF INVENTION: ENCODING THE LATTER; THEIR APPLICATION TO THE DETECTION OF
; TITLE OF INVENTION: CHAGAS DISEASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/138,736
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,917
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400
```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-138-736-12

Query Match      2.2%; Score 21; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 766 cgaagagaccatgaacactt 786
    |||||||||||||||||||
DB 1 CGAAGAGACCATGACCACTT 21

RESULT 10
US-09-332-782-18690
; Sequence 18690, Application US/09332782
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-754CON1
; CURRENT APPLICATION NUMBER: US/09/332,782
; CURRENT FILING DATE: 1999-06-14
; EARLIER APPLICATION NUMBER: US 09/181,430
; EARLIER FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 21027
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18690
; LENGTH: 391
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-332-782-18690

Query Match      2.0%; Score 20; DB 17; Length 391;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 740 ttctgaagtgtccctcgac 759
    |||||||||||||||||||
DB 353 ttctgaagtgtccctcgac 372

RESULT 11
US-09-515-694-18690
; Sequence 18690, Application US/09515694
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Blum, Linda
; APPLICANT: Cheung, Patrick
; APPLICANT: Damavandi, Simin
; APPLICANT: Drake, Jim
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Kotler, Janette
```

```

: APPLICANT: Labat, Ivan
: APPLICANT: Lee, Mon-Jae
: APPLICANT: Lomelli, Michelle
: APPLICANT: Nguyen, Hong
: APPLICANT: Nguyen, Linh
: APPLICANT: Nguyen, Lynne
: APPLICANT: Nguyen, Phuong
: APPLICANT: Nogra, Margie
: APPLICANT: Ojeda, Jesse
: APPLICANT: Palencia, Servando
: APPLICANT: Raisi, Fariba
: APPLICANT: Randhwa, Gurpreet
: APPLICANT: Sahourieh, Hannah
: APPLICANT: Sidhu, Navjwan
: APPLICANT: Smythe, Benjamin
: APPLICANT: Tkach, Joe
: APPLICANT: Tulpule, Mukul
: APPLICANT: Verna, Ron
: APPLICANT: Wachter, Adam
: APPLICANT: Wu, James
: APPLICANT: Yim, Kenneth
: TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
: FILE REFERENCE: 754CIP
: CURRENT APPLICATION NUMBER: US/09/515,694
: EARLIER APPLICATION NUMBER: 09/332,782
: EARLIER FILING DATE: 1999-06-14
: EARLIER APPLICATION NUMBER: 09/181,430
: EARLIER FILING DATE: 1998-10-28
: NUMBER OF SEQ ID NOS: 21027
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 18690
: LENGTH: 391
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-515-694-18690

Query Match      2.0%; Score 20; DB 19; Length 391;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 740 ttctgaagtgctccctgac 759
|||||
Db 353 ttctgaagtgctccctgac 372

RESULT 12
US-09-332-782-12269
: Sequence 12269, Application US/09332782
: GENERAL INFORMATION:
: APPLICANT: Hyseq, Inc.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
: FILE REFERENCE: 20411-754CON1
: CURRENT APPLICATION NUMBER: US/09/332,782
: CURRENT FILING DATE: 1999-06-14
: EARLIER APPLICATION NUMBER: 09/181,430
: EARLIER FILING DATE: 1998-10-28
: NUMBER OF SEQ ID NOS: 21027
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 12269
: LENGTH: 403
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-332-782-12269

Query Match      2.0%; Score 20; DB 17; Length 403;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy 740 ttctgaagtgctccctgac 759
|||||
Db 354 ttctgaagtgctccctgac 373

RESULT 13
US-09-515-694-12269
: Sequence 12269, Application US/09515694
: GENERAL INFORMATION:
: APPLICANT: Arterburn, Matthew
: APPLICANT: Asghari, Vida
: APPLICANT: Blum, Linda
: APPLICANT: Cheung, Patrick
: APPLICANT: Damavandi, Simin
: APPLICANT: Dickson, Mark
: APPLICANT: Drake, Jim
: APPLICANT: Drmanac, Radoje
: APPLICANT: Engleman, Carrie
: APPLICANT: Faulkner, Brandy
: APPLICANT: Fox, Melvin
: APPLICANT: Garcia, Veronica
: APPLICANT: Giedt, Gretchen
: APPLICANT: Jomek, Leni
: APPLICANT: Jones, Lee
: APPLICANT: Kita, David
: APPLICANT: Kofler, Janette
: APPLICANT: Labat, Ivan
: APPLICANT: Lee, Mon-Jae
: APPLICANT: Lomelli, Michelle
: APPLICANT: Nguyen, Hong
: APPLICANT: Nguyen, Linh
: APPLICANT: Nguyen, Lynne
: APPLICANT: Nguyen, Phuong
: APPLICANT: Nogra, Margie
: APPLICANT: Ojeda, Jesse
: APPLICANT: Palencia, Servando
: APPLICANT: Raisi, Fariba
: APPLICANT: Randhwa, Gurpreet
: APPLICANT: Sahourieh, Hannah
: APPLICANT: Sidhu, Navjwan
: APPLICANT: Smith, Benjamin
: APPLICANT: Smythe, Ashleigh
: APPLICANT: Tkach, Joe
: APPLICANT: Tulpule, Mukul
: APPLICANT: Verna, Ron
: APPLICANT: Wachter, Adam
: APPLICANT: Wu, James
: APPLICANT: Yim, Kenneth
: TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
: FILE REFERENCE: 754CIP
: CURRENT APPLICATION NUMBER: US/09/515,694
: CURRENT FILING DATE: 2000-02-29
: EARLIER APPLICATION NUMBER: 09/332,782
: EARLIER FILING DATE: 1999-06-14
: EARLIER APPLICATION NUMBER: 09/181,430
: EARLIER FILING DATE: 1998-10-28
: NUMBER OF SEQ ID NOS: 21027
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 12269
: LENGTH: 403
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-515-694-12269

Query Match      2.0%; Score 20; DB 19; Length 403;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 740 ttctgaagtgctccctgac 759
|||||
```

Db 354 ttctgaagttgctccctgac 373

RESULT 14

US-09-332-782-12260
; Sequence 12260, Application US/09332782
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-754CON1
; CURRENT APPLICATION NUMBER: US/09/332,782
; CURRENT FILING DATE: 1999-06-14
; EARLIER APPLICATION NUMBER: US 09/181,430
; EARLIER FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 21027
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12260
; LENGTH: 408
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(408)
; OTHER INFORMATION: n = A,T,C or G
US-09-332-782-12260

Query Match 2.0%; Score 20; DB 17; Length 408;

Best Local Similarity 100.0%; Pred. No. 15;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 740 ttctgaagttgctccctgac 759

Db 354 ttctgaagttgctccctgac 373

RESULT 15

US-09-515-694-12260
; Sequence 12260, Application US/09515694
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Blum, Linda
; APPLICANT: Cheung, Patrick
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandt
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jomek, Leol
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Kofler, Janette
; APPLICANT: Labat, Ivan
; APPLICANT: Lee, Won-Jae
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nguyen, Hong
; APPLICANT: Nguyen, Linh
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Ojeda, Jesse
; APPLICANT: Palencia, Servando
; APPLICANT: Ralski, Farida
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sabourieh, Hannah
; APPLICANT: Sidhu, Navjivan
; APPLICANT: Smith, Benjamin

; APPLICANT: Smythe, Ashleigh

; APPLICANT: Tkach, Joe

; APPLICANT: Tulpule, Mukul

; APPLICANT: Verna, Ron

; APPLICANT: Wachter, Adam

; APPLICANT: Wu, James

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

; FILE REFERENCE: 754CIP

; CURRENT APPLICATION NUMBER: US/09/515,694

; CURRENT FILING DATE: 2000-02-29

; EARLIER APPLICATION NUMBER: 09/332,782

; EARLIER FILING DATE: 1999-06-14

; EARLIER APPLICATION NUMBER: 09/181,430

; EARLIER FILING DATE: 1998-10-28

; NUMBER OF SEQ ID NOS: 21027

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 12260

; LENGTH: 408

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(408)

; OTHER INFORMATION: n = A,T,C or G
US-09-515-694-12260

Query Match 2.0%; Score 20; DB 19; Length 408;

Best Local Similarity 100.0%; Pred. No. 15;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 740 ttctgaagttgctccctgac 759

Db 354 ttctgaagttgctccctgac 373

RESULT 16

US-09-654-617-126156
; Sequence 126156, Application US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 126156
; LENGTH: 718
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-654-617-126156

Query Match 2.0%; Score 20; DB 25; Length 718;

Best Local Similarity 100.0%; Pred. No. 15;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 212 ctctgcgcgtgctgac 231

Db 586 ctctgcgcgtgctgac 605

RESULT 17

US-09-684-016-126156
; Sequence 126156, Application US/09684016
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D

; CURRENT APPLICATION NUMBER: US/09/684,016
; CURRENT FILING DATE: 2000-10-10 US 09/654,617
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 126156
; LENGTH: 718
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-684-016-126156

Query Match 2.0%; Score 20; DB 27; Length 718;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 212 ctctgcgcgtcgtcatcc 231
|||||
Db 586 ctctgcgcgtcgtcatcc 605

RESULT 18
US-09-620-392-64793/c
; Sequence 64793, Application US/09620392
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: McIninch, James
; TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof
; FILE REFERENCE: 38-21(51237)E
; CURRENT APPLICATION NUMBER: US/09/620,392
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 69652
; SEQ ID NO 64793
; LENGTH: 5679
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-620-392-64793

Query Match 2.0%; Score 20; DB 24; Length 5679;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 215 ctgcgcgtcgtcatccct 234
|||||
Db 3043 CTGCCGCTGTCGATCCCT 3024

RESULT 19
US-09-702-134-21010/c
; Sequence 21010, Application US/09702134
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: McIninch, James
; APPLICANT: Mu, Wei
; TITLE OF INVENTION: Plant Genome Sequence and Uses Thereof
; FILE REFERENCE: 38-21(51237)F
; CURRENT APPLICATION NUMBER: US/09/702,134
; CURRENT FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 52202
; SEQ ID NO 21010
; LENGTH: 5679
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-702-134-21010

Query Match 2.0%; Score 20; DB 28; Length 5679;

Best Local Similarity 100.0%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 215 ctgcgcgtcgtcatccct 234
|||||
Db 3043 CTGCCGCTGTCGATCCCT 3024

RESULT 20
US-60-173-464-11355
; Sequence 11355, Application US/60173464
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W.D.
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00173
; CURRENT APPLICATION NUMBER: US/60/173,464
; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 30269
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11355
; LENGTH: 11355
; TYPE: DNA
; ORGANISM: Drosophila
US-60-173-464-11355

Query Match 2.0%; Score 20; DB 49; Length 11335;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 50 ggcctcgtgctgcatgctg 69
|||||
Db 6835 ggcctcgtgctgcatgctg 6854

RESULT 21
US-60-191-637-13953/c
; Sequence 13953, Application US/60191637
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO00392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13953
; LENGTH: 11335
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-60-191-637-13953

Query Match 2.0%; Score 20; DB 51; Length 11335;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 50 ggcctcgtgctgcatgctg 69
|||||
Db 4501 GGCCTCTGCTGCTGCTGCTG 4482

RESULT 22
US-60-191-681-10986/c
; Sequence 10986, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND

;; TITLE OF INVENTION: USES THEREOF.
;; FILE REFERENCE: C1000390
;; CURRENT APPLICATION NUMBER: US/60/191,681
;; CURRENT FILING DATE: 2000-03-23
;; NUMBER OF SEQ ID NOS: 30973
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 10986
;; LENGTH: 11335
;; TYPE: DNA
;; ORGANISM: DROSOPHILA
US-60-191-681-10986

Query Match 2.0%; Score 20; DB 51; Length 11335;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 ggcgtcgtgctgacatgctg 69
|||||
DB 4501 GCGTCTGCGCTGACATGCTG 4482

RESULT 23
US-60-167-217-13998
;; Sequence 13998, Application US/60167217
;; GENERAL INFORMATION:
;; APPLICANT: LI, Peter W. D.
;; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
;; TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES
;; TITLE OF INVENTION: THEREOF
;; FILE REFERENCE: CL000152
;; CURRENT APPLICATION NUMBER: US/60/167,217
;; CURRENT FILING DATE: 1999-11-24
;; NUMBER OF SEQ ID NOS: 23195
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 13998
;; LENGTH: 11338
;; TYPE: DNA
;; ORGANISM: Drosophila
US-60-167-217-13998

Query Match 2.0%; Score 20; DB 48; Length 11338;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 ggcgtcgtgctgacatgctg 69
|||||
DB 6838 ggcgtcgtgctgacatgctg 6857

RESULT 24
US-09-534-859-19/C
;; Sequence 19, Application US/09534859
;; GENERAL INFORMATION:
;; APPLICANT: Bush, David F.
;; APPLICANT: Last, Robert L.
;; APPLICANT: Levin, Irena M.
;; APPLICANT: Norris, Susan R.
;; APPLICANT: Parnell, Laurence D.
;; APPLICANT: Rounsley, Steven D.
;; APPLICANT: Wiegand, Roger C.
;; TITLE OF INVENTION: PLANT POLYMORPHIC MARKERS AND USES THEREOF
;; FILE REFERENCE: 38-10(15493)B
;; CURRENT APPLICATION NUMBER: US/09/534,859
;; CURRENT FILING DATE: 2000-03-29
;; NUMBER OF SEQ ID NOS: 1127
;; SEQ ID NO 19
;; LENGTH: 91470
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
US-09-534-859-19

Query Match 2.0%; Score 20; DB 20; Length 91470;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 ctctcgcgctgctgcatcc 231
|||||
DB 42747 CTCTCGCGCTGCTGCATCC 42728

RESULT 25
US-09-528-237A-1285
;; Sequence 1285, Application US/09528237A
;; GENERAL INFORMATION:
;; APPLICANT: Venter, J. Craig
;; TITLE OF INVENTION: The Drosophila Genome, Primary Nucleic
;; TITLE OF INVENTION: Acid Sequences, Systems Containing The Nucleic Acid
;; TITLE OF INVENTION: Sequences and Uses Thereof
;; FILE REFERENCE: CL000284
;; CURRENT APPLICATION NUMBER: US/09/528,237A
;; CURRENT FILING DATE: 2000-03-17
;; NUMBER OF SEQ ID NOS: 2926
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1285
;; LENGTH: 226706
;; TYPE: DNA
;; ORGANISM: Drosophila
US-09-528-237A-1285

Query Match 2.0%; Score 20; DB 19; Length 226706;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 ggcgtcgtgctgacatgctg 69
|||||
DB 213597 ggcgtcgtgctgacatgctg 213616

RESULT 26
US-60-182-316-2055/C
;; Sequence 2055, Application US/60182316
;; GENERAL INFORMATION:
;; APPLICANT: Curtis, Anne L.
;; APPLICANT: Lagace, Robert E.
;; APPLICANT: Klingler, Tod M.
;; APPLICANT: Stuve, Laura L.
;; TITLE OF INVENTION: CpG Island Polynucleotides
;; FILE REFERENCE: PX-0003 P
;; CURRENT APPLICATION NUMBER: US/60/182,316
;; CURRENT FILING DATE: 2000-02-10
;; NUMBER OF SEQ ID NOS: 14,630
;; SOFTWARE: PERL Program
;; SEQ ID NO 2055
;; LENGTH: 718
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc.feature
;; OTHER INFORMATION: Incyte ID No: CPG_991027_B15_masked_fa.Contig49773
US-60-182-316-2055

Query Match 1.9%; Score 19; DB 50; Length 718;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 gccgcgagcgatcgctg 298
|||||
DB 213 GCCGCGCGCGATCGCTG 195

RESULT 27

```
US-09-620-392-32691
; Sequence 32691, Application US/09620392
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof
; FILE REFERENCE: 38-21(51237)E
; CURRENT APPLICATION NUMBER: US/09/620,392
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 69652
; SEQ ID NO 32691
; LENGTH: 9855
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-620-392-32691

Query Match
Best Local Similarity 1.9%; Score 19; DB 24; Length 9855;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 gccggccacgacgagctct 405
    ||||||||||||||||
Db 259 gccggccacgacgagctct 277

RESULT 28
US-09-620-392-67713
; Sequence 67713, Application US/09620392
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof
; FILE REFERENCE: 38-21(51237)E
; CURRENT APPLICATION NUMBER: US/09/620,392
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 69652
; SEQ ID NO 67713
; LENGTH: 9895
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-620-392-67713

Query Match
Best Local Similarity 1.9%; Score 19; DB 24; Length 9895;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 gccggccacgacgagctct 405
    ||||||||||||||||
Db 8100 gccggccacgacgagctct 8118

RESULT 29
US-09-407-562-32/C
; Sequence 32, Application US/09407562
; GENERAL INFORMATION:
; APPLICANT: Katheryn Meek
; TITLE OF INVENTION: Genetic Test For Equine Severe
; TITLE OF INVENTION: Combined Immunodeficiency Disease
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Benjamin A. Adler
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/407,562
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/970,269
; FILING DATE: November 14, 1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler Ph.D., Benjamin A.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5860
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-777-2321
; TELEFAX: 713-777-6908
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11873 bp
; TYPE: nucleic acid
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: other nucleic acid
; HYPOTHEICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; FEATURE:
US-09-407-562-32
```

```
Query Match
Best Local Similarity 1.9%; Score 19; DB 18; Length 11873;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 actctgcgcgtctgcgat 229
    ||||||||||||||||
Db 6446 actctgcgcgtctgcgat 6428

RESULT 30
US-09-407-562-31/C
; Sequence 31, Application US/09407562
; GENERAL INFORMATION:
; APPLICANT: Katheryn Meek
; TITLE OF INVENTION: Genetic Test For Equine Severe
; TITLE OF INVENTION: Combined Immunodeficiency Disease
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Benjamin A. Adler
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/407,562
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/970,269
; FILING DATE: November 14, 1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler Ph.D., Benjamin A.
```

```

:
:   REGISTRATION NUMBER: 35,423
:   REFERENCE/DOCKET NUMBER: D5860
:   TELECOMMUNICATION INFORMATION:
:   TELEPHONE: 713-777-2321
:   TELEFAX: 713-777-6908
:   INFORMATION FOR SEQ ID NO: 31:
:   SEQUENCE CHARACTERISTICS:
:   LENGTH: 11878 bp
:   TYPE: nucleic acid
:   STRANDEDNESS: double stranded
:   TOPOLOGY: linear
:   MOLECULE TYPE:
:   DESCRIPTION: other nucleic acid
:   HYPOTHETICAL: no
:   ANTI-SENSE: no
:   ORIGINAL SOURCE:
:   FEATURE:
:   US-09-407-562-31

```

```

Query Match      1.9%; Score 19; DB 18; Length 11878;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 211 acctctgcgcgtcgtcat 229
    |||
DB 6446 ACTCTGCCGCTGCTGCAT 6428

```

```

RESULT 31
US-09-407-562-28/C
: Sequence 28, Application US/09407562
:
:   GENERAL INFORMATION:
:   APPLICANT: Kathryn Meek
:   TITLE OF INVENTION: Genetic Test For Equine Severe
:   TITLE OF INVENTION: Combined Immunodeficiency Disease
:   NUMBER OF SEQUENCES: 32
:   CORRESPONDENCE ADDRESS:
:   ADDRESSEE: Dr. Benjamin A. Adler
:   STREET: 8011 Candle Lane
:   CITY: Houston
:   STATE: Texas
:   COUNTRY: USA
:   ZIP: 77071
:   COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   COMPUTER: Apple
:   OPERATING SYSTEM: Macintosh
:   SOFTWARE: Microsoft Word for Macintosh
:   CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/09/407,562
:   FILING DATE:
:   CLASSIFICATION:
:   PRIOR APPLICATION DATA:
:   APPLICATION NUMBER: 08/970,269
:   FILING DATE: November 14, 1997
:   CLASSIFICATION:
:   ATTORNEY/AGENT INFORMATION:
:   NAME: Adler Ph.D., Benjamin A.
:   REGISTRATION NUMBER: 35,423
:   REFERENCE/DOCKET NUMBER: D5860
:   TELECOMMUNICATION INFORMATION:
:   TELEPHONE: 713-777-2321
:   TELEFAX: 713-777-6908
:   INFORMATION FOR SEQ ID NO: 28:
:   SEQUENCE CHARACTERISTICS:
:   LENGTH: 11883 bp
:   TYPE: nucleic acid
:   STRANDEDNESS: double stranded
:   TOPOLOGY: linear
:   MOLECULE TYPE:
:   DESCRIPTION: other nucleic acid
:   HYPOTHETICAL: no

```

```

:   ANTI-SENSE: no
:   ORIGINAL SOURCE:
:   US-09-407-562-28

```

```

Query Match      1.9%; Score 19; DB 18; Length 11883;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 211 acctctgcgcgtcgtcat 229
    |||
DB 6449 ACTTCTGCCGCTGCTGCAT 6431

```

```

RESULT 32
US-09-620-392-10632/C
: Sequence 10632, Application US/09620392
:
:   GENERAL INFORMATION:
:   APPLICANT: Boukharov, Andrey A.
:   APPLICANT: Kovalic, David K.
:   APPLICANT: Liu, Jindong
:   APPLICANT: McIninch, James
:   TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof
:   FILE REFERENCE: 38-21(51237)E
:   CURRENT APPLICATION NUMBER: US/09/620,392
:   CURRENT FILING DATE: 2000-07-19
:   NUMBER OF SEQ ID NOS: 69652
:   SEQ ID NO 10632
:   LENGTH: 19505
:   TYPE: DNA
:   ORGANISM: Oryza sativa
:   US-09-620-392-10632

```

```

Query Match      1.9%; Score 19; DB 24; Length 19505;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 387 gccggccacgacgaggtct 405
    |||
DB 19008 GCCGGCCACGACGAGGTCT 18990

```

```

RESULT 33
US-09-620-392-4251/C
: Sequence 4251, Application US/09620392
:
:   GENERAL INFORMATION:
:   APPLICANT: Boukharov, Andrey A.
:   APPLICANT: Kovalic, David K.
:   APPLICANT: Liu, Jindong
:   APPLICANT: McIninch, James
:   TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof
:   FILE REFERENCE: 38-21(51237)E
:   CURRENT APPLICATION NUMBER: US/09/620,392
:   CURRENT FILING DATE: 2000-07-19
:   NUMBER OF SEQ ID NOS: 69652
:   SEQ ID NO 4251
:   LENGTH: 25670
:   TYPE: DNA
:   ORGANISM: Oryza sativa
:   OTHER INFORMATION: unsure at all n locations
:   US-09-620-392-4251

```

```

Query Match      1.9%; Score 19; DB 24; Length 25670;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 387 gccggccacgacgaggtct 405
    |||
DB 3582 GCCGGCCACGACGAGGTCT 3564

```



```
RESULT 34
US-09-702-134-5295/c
: Sequence 5295, Application US/09702134
: GENERAL INFORMATION:
: APPLICANT: Boukharov, Andrey A.
: APPLICANT: Cao, Yongwei
: APPLICANT: Kovalic, David K.
: APPLICANT: Liu, Jindong
: APPLICANT: McIninch, James
: APPLICANT: Wu, Mei
: TITLE OF INVENTION: Plant Genome Sequence and Uses Thereof
: FILE REFERENCE: 38-21(51237)F
: CURRENT APPLICATION NUMBER: US/09/702.134
: CURRENT FILING DATE: 2000-10-31
: NUMBER OF SEQ ID NOS: 52202
: SEQ ID NO 5295
: LENGTH: 31556
: TYPE: DNA
: ORGANISM: Oryza sativa
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (1)..(31556)
: OTHER INFORMATION: unsure at all n locations
US-09-702-134-5295

Query Match 1.9%; Score 19; DB 28; Length 31556;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 387 gccggccagcagaggtct 405
|||||
Db 9468 GCCGCCACGACGAGGTCT 9450

RESULT 35
US-09-702-134-5301
: Sequence 5301, Application US/09702134
: GENERAL INFORMATION:
: APPLICANT: Boukharov, Andrey A.
: APPLICANT: Cao, Yongwei
: APPLICANT: Kovalic, David K.
: APPLICANT: Liu, Jindong
: APPLICANT: McIninch, James
: APPLICANT: Wu, Mei
: TITLE OF INVENTION: Plant Genome Sequence and Uses Thereof
: FILE REFERENCE: 38-21(51237)F
: CURRENT APPLICATION NUMBER: US/09/702.134
: CURRENT FILING DATE: 2000-10-31
: NUMBER OF SEQ ID NOS: 52202
: SEQ ID NO 5301
: LENGTH: 159989
: TYPE: DNA
: ORGANISM: Oryza sativa
US-09-702-134-5301

Query Match 1.9%; Score 19; DB 28; Length 159989;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 387 gccggccagcagaggtct 405
|||||
Db 32451 gccggccagcagaggtct 32469

RESULT 36
US-08-988-242-4/c
: Sequence 4, Application US/08988242
: GENERAL INFORMATION:
: APPLICANT: PARANHOS-BACCALA, GLAUCIA
: APPLICANT: LESENECHAL, MYLENE
: APPLICANT: JOLIVET, MICHEL
```

```
APPLICANT: MANDRAND, BERNARD
: TITLE OF INVENTION: NEW TRYPANOSOMA CRUZI ANTIGEN, GENE
: TITLE OF INVENTION: ENCODING THEREFOR, AND METHODS OF DETECTING AND TREATING
: TITLE OF INVENTION: CHAGAS DISEASE
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OLIFF & BERRIDGE, PLC
: STREET: P.O. BOX 19928
: CITY: Alexandria
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22314
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/988.242
: FILING DATE: 10-DEC-1997
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Berridge, William P.
: REGISTRATION NUMBER: 30,024
: REFERENCE/DOCKET NUMBER: MPB 36400A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-836-6400
: TELEFAX: 703-836-2787
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 18 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
US-08-988-242-4

Query Match 1.8%; Score 18; DB 13; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 211 actctgcgcgcgtcgtca 228
|||||
Db 18 ACTTCTGCCGCTGCTGCA 1

RESULT 37
US-09-138-736-5
: Sequence 5, Application US/09138736
: GENERAL INFORMATION:
: APPLICANT: PARANHOS-BACCALA, GLAUCIA
: APPLICANT: LESENECHAL, MYLENE
: APPLICANT: JOLIVET, MICHEL
: TITLE OF INVENTION: NEW TRYPANOSOMA CRUZI ANTIGEN, AND GENE
: TITLE OF INVENTION: ENCODING THE LATTER; THEIR APPLICATION TO THE DETECTION OF
: TITLE OF INVENTION: CHAGAS DISEASE
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Oliff & Berridge
: STREET: 700 South Washington Street, Suite 300
: CITY: Alexandria
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22314
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/138.736
: FILING DATE:
```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/480,917
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 36400
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-138-736-5

Query Match 1.8%; Score 18; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 tcggcgactgacgcgcgcg 189
|||||
DB 1 TCGGCGACTGACCGCGCG 18

RESULT 38
US-09-138-736-8/c
Sequence 8, Application US/09138736
GENERAL INFORMATION:
APPLICANT: PARANHOS-BACCALA, Glaucia
APPLICANT: LESENECHAL, Mylene
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: NEW TRYPAOSOMA CRUZI ANTIGEN, AND GENE
TITLE OF INVENTION: ENCODING THE LATTER; THEIR APPLICATION TO THE DETECTION OF
TITLE OF INVENTION: CHAGAS DISEASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Oliff & Berridge
STREET: 700 South Washington Street, Suite 300
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/138,736
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/480,917
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 36400
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
US-09-138-736-8

Query Match 1.8%; Score 18; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 actcttcgcgcgtgcatcca 228
|||||
DB 18 ACTTCGTGCCGCTGCTGCA 1

RESULT 39
US-09-366-691A-1504/c
Sequence 1504, Application US/09366691A
GENERAL INFORMATION:
APPLICANT: Rosen, Craig A., et al.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products. 38
FILE REFERENCE: PO-38
CURRENT APPLICATION NUMBER: US/09/366,691A
CURRENT FILING DATE: 1999-08-04
PRIOR APPLICATION NUMBER: 60/095,485
PRIOR FILING DATE: 1998-08-05
NUMBER OF SEQ ID NOS: 8319
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1504
LENGTH: 70
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (10)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (60)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (65)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (70)
OTHER INFORMATION: n equals a,t,g, or c
US-09-366-691A-1504

Query Match 1.8%; Score 18; DB 17; Length 70;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 tcgcgcgtgcatccccc 233
|||||
DB 31 TGGCGCTGCTGATCCGCC 14

RESULT 40
US-09-076-667-928/c
Sequence 928, Application US/09076667
GENERAL INFORMATION:
APPLICANT: Gooding, Douglas H.
APPLICANT: Stuve, Laura L.
APPLICANT: Stuart, Susan G.
APPLICANT: Ito, Laura Y.
APPLICANT: Akerbloom, Ingrid E.
APPLICANT: Deleageane, Angelo M.
APPLICANT: Naughton, Rebecca E.
APPLICANT: Klingler, Tod M.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
TITLE OF INVENTION: HUMAN BRAIN
NUMBER OF SEQUENCES: 4483
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE

CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/076,667
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PD-0370P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 928:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
IMMEDIATE SOURCE:
CLONE: 3244053H1
US-09-076-667-928

Query Match 1.8%; Score 18; DB 14; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 111 agaacagcagccagcca 128
|||||
DB 29 AGAACAGCAGCCAGGCCA 12

RESULT 41
US-09-540-229-15688/C
Sequence 15688, Application US/09540229
GENERAL INFORMATION:
APPLICANT: Seilhamer, Jeffrey J.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullaby, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES OF NERVOUS SYSTEM AND SENSORY ORGANS
FILE REFERENCE: PD-1033 CIP
CURRENT APPLICATION NUMBER: US/09/540,229
CURRENT FILING DATE: 2000-03-31
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 193582
SOFTWARE: PERL Program
SEQ ID NO 15688
LENGTH: 85
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
OTHER INFORMATION: Incyte ID No: hu00290743
NAME/KEY: unsure
LOCATION: 3, 81
OTHER INFORMATION: a, t, c, g, or other
US-09-540-229-15688

Query Match 1.8%; Score 18; DB 21; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 111 agaacagcagccagcca 128
|||||
DB 29 AGAACAGCAGCCAGGCCA 12

RESULT 42
US-60-048-002-928/C
Sequence 928, Application US/60048002
GENERAL INFORMATION:
APPLICANT: Gooding, Douglas H.
APPLICANT: Stuve, Laura L.
APPLICANT: Stuart, Susan G.
APPLICANT: Ito, Laura Y.
APPLICANT: Akerblom, Ingrid E.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Naughton, Rebecca E.
APPLICANT: Klingler, Tod M.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
TITLE OF INVENTION: HUMAN BRAIN
NUMBER OF SEQUENCES: 4483
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/048,002
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PD-0370P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 928:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
IMMEDIATE SOURCE:
CLONE: 3244053H1
US-60-048-002-928

Query Match 1.8%; Score 18; DB 36; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 agaacagcagccagcca 128
|||||
DB 29 AGAACAGCAGCCAGGCCA 12

RESULT 43
US-09-513-991-2039
Sequence 2039, Application US/09513991
GENERAL INFORMATION:
APPLICANT: Gearing, David P.
APPLICANT: Fraser, Christopher C.
APPLICANT: Holtzman, Douglas A.

```

; APPLICANT: Jakubowski, Joseph A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM
; FILE REFERENCE: 1600.1083-001
; CURRENT APPLICATION NUMBER: US/09/513,991
; EARLIER FILING DATE: 2000-03-08
; EARLIER APPLICATION NUMBER: 60/123,397
; NUMBER OF SEQ ID NOS: 3378
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2039
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(180)
; OTHER INFORMATION: n = A,T,C or G
US-09-513-991-2039

```

```

Query Match
Best Local Similarity 1.8%; Score 18; DB 19; Length 180;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 92 tccggtctctccgcaggg 109
|||||
DB 1 tccggtctctccgcaggg 18

```

```

RESULT 44
US-09-654-617-141143/c
; Sequence 141143, Application US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 141143
; LENGTH: 194
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-654-617-141143

```

```

Query Match
Best Local Similarity 1.8%; Score 18; DB 25; Length 194;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 930 caacatcacgaagggt 947
|||||
DB 178 CAACATCAGCAGAGGGGT 161

```

```

RESULT 45
US-09-684-016-141143/c
; Sequence 141143, Application US/09684016
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/684,016
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/654,617
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 141143
; LENGTH: 194
; TYPE: DNA

```

```

; ORGANISM: Arabidopsis thaliana
US-09-684-016-141143

```

```

Query Match
Best Local Similarity 1.8%; Score 18; DB 27; Length 194;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 930 caacatcacgaagggt 947
|||||
DB 178 CAACATCAGCAGAGGGGT 161

```

```

RESULT 46
US-09-654-617-162528/c
; Sequence 162528, Application US/09654617
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 162528
; LENGTH: 196
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana columbia
US-09-654-617-162528

```

```

Query Match
Best Local Similarity 1.8%; Score 18; DB 25; Length 196;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 930 caacatcacgaagggt 947
|||||
DB 153 CAACATCAGCAGAGGGGT 136

```

```

RESULT 47
US-09-684-016-162528/c
; Sequence 162528, Application US/09684016
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/684,016
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/654,617
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 162528
; LENGTH: 196
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana columbia
US-09-684-016-162528

```

```

Query Match
Best Local Similarity 1.8%; Score 18; DB 27; Length 196;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 930 caacatcacgaagggt 947
|||||
DB 153 CAACATCAGCAGAGGGGT 136

```

```

RESULT 48
US-09-654-617-141154/c
; Sequence 141154, Application US/09654617
; GENERAL INFORMATION:

```

```
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 141154
; LENGTH: 217
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-654-617-141154

Query Match
Best Local Similarity 1.8%; Score 18; DB 25; Length 217;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 930 caacatcagaaggggt 947
|||||
DB 168 CAACATCAGGAAGGGGT 151

RESULT 49
US-09-684-016-141154/c
; Sequence 141154, Application US/09684016
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/684,016
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/654,617
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 141154
; LENGTH: 217
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-684-016-141154

Query Match
Best Local Similarity 1.8%; Score 18; DB 27; Length 217;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 930 caacatcagaaggggt 947
|||||
DB 168 CAACATCAGGAAGGGGT 151

RESULT 50
US-08-810-326-2151/c
; Sequence 2151, Application US/08810326
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Kingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; NUMBER OF SEQUENCES: 3314
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
```

```
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,326
; FILING DATE: HERWITH
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/012,699
; FILING DATE: FEBRUARY 29, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/015,173
; FILING DATE: APRIL 10, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 2151:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 765907
US-08-810-326-2151
```

```
Query Match
Best Local Similarity 1.8%; Score 18; DB 12; Length 234;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 gtagctgcgtcctgctg 61
|||||
DB 136 GTAGCTGCCTCCTGCTG 119
```

Search completed: September 21, 2001, 16:22:20
Job time: 9685 sec

```
975 15 1.5 573 8 US-60-253-652-15676 Sequence 15676, A
976 15 1.5 573 8 US-60-253-652-19100 Sequence 19100, A
977 15 1.5 575 7 US-09-834-366-12963 Sequence 12963, A
978 15 1.5 577 8 US-60-253-652-6296 Sequence 6296, Ap
979 15 1.5 579 6 US-09-810-173-52 Sequence 52, Appl
980 15 1.5 582 8 US-60-253-457-14243 Sequence 14243, A
981 15 1.5 583 7 US-09-838-601-221 Sequence 221, App
982 15 1.5 584 8 US-60-252-833-29966 Sequence 29966, A
983 15 1.5 586 6 US-60-253-652-16456 Sequence 16456, A
984 15 1.5 587 6 US-09-795-300-188 Sequence 188, App
985 15 1.5 588 8 US-60-252-833-28454 Sequence 28454, A
986 15 1.5 588 8 US-60-252-833-30247 Sequence 30247, A
987 15 1.5 589 8 US-60-252-833-42071 Sequence 42071, A
988 15 1.5 591 8 US-60-255-619-17807 Sequence 17807, A
989 15 1.5 594 7 US-09-825-422-30 Sequence 30, Appl
990 15 1.5 595 8 US-60-253-378-2609 Sequence 2609, Ap
991 15 1.5 596 6 US-09-898-888-10715 Sequence 10715, A
992 15 1.5 598 6 US-09-898-888-22239 Sequence 22239, A
993 15 1.5 598 6 US-09-758-440-122 Sequence 122, Appl
994 15 1.5 599 1 PCT-US01-15332-54 Sequence 122, Appl
995 15 1.5 599 1 PCT-US01-08631-16760 Sequence 16760, A
996 15 1.5 599 8 US-60-253-456-17444 Sequence 17444, A
997 15 1.5 599 8 US-60-253-653-10667 Sequence 10667, A
998 15 1.5 600 7 US-09-825-422-243 Sequence 243, App
999 15 1.5 601 8 US-60-278-232-11399 Sequence 11399, A
c1000 15 1.5 602 8 US-60-253-652-19417 Sequence 19417, A
```

ALIGNMENTS

```
RESULT 1
US-09-138-735-1
; Sequence 1, Application US/09138735
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: LESENECHAL, Mylene
; TITLE OF INVENTION: TRYPAOSOMA CRUZI ANTIGEN, GENE ENCODING THEREFOR AND METHODS OF
; FILE REFERENCE: MPB 36400B
; CURRENT APPLICATION NUMBER: US/09/138, 735
; CURRENT FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: US 08/480, 917
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: FR 94/10132
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 1
; LENGTH: 3402
; TYPE: DNA
; ORGANISM: Trypanosoma cruzi
US-09-138-735-1
```

```
Query Match 100.0%; Score 976; DB 6; Length 3402;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 976; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 caagtcagtgatcagcgtcttcttcaatcgtacacgcagcgtacgtctcgtcgtc 60
D 1232 caagtcagtgatcagcgtcttcttcaatcgtacacgcagcgtacgtctcgtcgtc 1291
QY 61 gacatgcgtatcgtatgacatcttcacatccgcgtcccccgcagcaggaagacacag 120
D 1292 gacatgcgtatcgtatgacatcttcacatccgcgtcccccgcagcaggaagacacag 1351
QY 121 ccagggccaaaacatcgtgtagtgcgacgcggaacccgggggtgtgtcctcgggac 180
D 1352 ccagggccaaaacatcgtgtagtgcgacgcggaacccgggggtgtgtcctcgggac 1411
QY 181 gaacggcgagtagcagtcataacacatcgtcgcgtcgtcgtacatccctcatca 240
```

```
|||||
D 1412 gaacggcgagtagcagtcataacacatcgtcgcgtcgtcgtacatccctcatca 1471
QY 241 cccctcttcaacgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 300
D 1472 cccctcttcaacgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1531
QY 301 ccgcacgtggggagcaagacatctgtaactagtagtaactagtagtagtaactagtagta 360
D 1532 ccgcacgtggggagcaagacatctgtaactagtagtagtagtagtagtagtagtagtag 1591
QY 361 caaagagcgtcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 420
D 1592 caaagagcgtcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1651
QY 421 acgactaccccccgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 480
D 1652 acgactaccccccgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1711
QY 481 ggaatggtgagcagtaatagttggtggtcagcagcagcagcagcagcagcagcagcag 540
D 1712 ggaatggtgagcagtaatagttggtggtcagcagcagcagcagcagcagcagcagcag 1771
QY 541 ccaacagaggaagaaagagcgtccgcagcagcagcagcagcagcagcagcagcagcagc 600
D 1772 ccaacagaggaagaaagagcgtccgcagcagcagcagcagcagcagcagcagcagcagc 1831
QY 601 acacgcccgcgtcttcaacagcgtctctccgcagcagcagcagcagcagcagcagcag 660
D 1832 acacgcccgcgtcttcaacagcgtctctccgcagcagcagcagcagcagcagcagcag 1891
QY 661 gcagcagatatacagtcgagcagcagcagcagcagcagcagcagcagcagcagcagcag 720
D 1892 gcagcagatatacagtcgagcagcagcagcagcagcagcagcagcagcagcagcagcag 1951
QY 721 ataaccacagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 780
D 1952 ataaccacagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 2011
QY 781 caactctgacatcgtggtttagagcagcagcagcagcagcagcagcagcagcagcagcag 840
D 2012 caactctgacatcgtggtttagagcagcagcagcagcagcagcagcagcagcagcagcag 2071
QY 841 ccgcaaacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 900
D 2072 ccgcaaacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2131
QY 901 ctgtctcagtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 960
D 2132 ctgtctcagtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2191
QY 961 aacgagccatattgt 976
D 2192 aacgagccatattgt 2207
```

```
RESULT 2
US-09-138-735-9
; Sequence 9, Application US/09138735
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: LESENECHAL, Mylene
; TITLE OF INVENTION: TRYPAOSOMA CRUZI ANTIGEN, GENE ENCODING THEREFOR AND METHODS.
; FILE REFERENCE: MPB 36400B
; CURRENT APPLICATION NUMBER: US/09/138, 735
; CURRENT FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: US 08/480, 917
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: FR 94/10132
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 13
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: probe/primer
US-09-138-735-9
```

```
Query Match          2.3%; Score 22; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 35 cagcgacgtagctgctcct 56
Db 1 cagcgacgtagctgctcct 22
```

```
RESULT 3
US-09-138-735-10/c
; Sequence 10, Application US/09138735
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACALA, Glaucia
; APPLICANT: LESENECHAL, Mylene
; APPLICANT: JOLIVET, Michel
; TITLE OF INVENTION: TRYPANOSOMA CRUZI ANTIGEN, GENE ENCODING THEREFOR AND METHODS OF
; TITLE OF INVENTION: DETECTING AND TREATING CHAGAS DISEASE
; FILE REFERENCE: WPB 36400B
; CURRENT APPLICATION NUMBER: US/09/138,735
; CURRENT FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: US 08/480,917
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: FR 94/10132
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: probe/primer
US-09-138-735-10
```

```
Query Match          2.2%; Score 21; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 956 gtgtgaagagccattatgt 976
Db 21 gtgtgaagagccattatgt 1
```

```
RESULT 4
US-09-138-735-12
; Sequence 12, Application US/09138735
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACALA, Glaucia
; APPLICANT: LESENECHAL, Mylene
; APPLICANT: JOLIVET, Michel
; TITLE OF INVENTION: TRYPANOSOMA CRUZI ANTIGEN, GENE ENCODING THEREFOR AND METHODS OF
; TITLE OF INVENTION: DETECTING AND TREATING CHAGAS DISEASE
; FILE REFERENCE: WPB 36400B
; CURRENT APPLICATION NUMBER: US/09/138,735
; CURRENT FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: US 08/480,917
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: FR 94/10132
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 12
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: probe/primer
US-09-138-735-12
```

```
Query Match          2.2%; Score 21; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 766 cgaagagaccatgaacaact 786
Db 1 cgaagagaccatgaacaact 21
```

```
RESULT 5
US-09-803-736-19/c
; Sequence 19, Application US/09803736
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Levin, Irena M.
; APPLICANT: Norris, Susan R.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)D
; CURRENT APPLICATION NUMBER: US/09/803,736
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/534,859
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: Identified by Attorney Docket number 04983.0206CPUS01 38
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 1582
; SEQ ID NO 19
; LENGTH: 91470
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-803-736-19
```

```
Query Match          2.0%; Score 20; DB 6; Length 91470;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 212 ctctgcgcgtctgcatcc 231
Db 42747 ctctgcgcgcgtctgcatcc 42728
```

```
RESULT 6
US-09-138-735-5
; Sequence 5, Application US/09138735
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACALA, Glaucia
; APPLICANT: LESENECHAL, Mylene
; APPLICANT: JOLIVET, Michel
; TITLE OF INVENTION: TRYPANOSOMA CRUZI ANTIGEN, GENE ENCODING THEREFOR AND METHODS
; TITLE OF INVENTION: DETECTING AND TREATING CHAGAS DISEASE
; FILE REFERENCE: WPB 36400B
; CURRENT APPLICATION NUMBER: US/09/138,735
; CURRENT FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: US 08/480,917
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: FR 94/10132
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 18
; TYPE: DNA
```

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: primer
US-09-138-735-5

Query Match 1.8%; Score 18; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 172 tcgggacactgacgcgcgcg 189
|||||
Db 1 tcgggacactgacgcgcgcg 18

RESULT 7
US-09-138-735-8/c
; Sequence 8, Application US/09138735
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: LESENECHAL, Mylene
; APPLICANT: JOLIVET, Michel
; TITLE OF INVENTION: TRYPAOSOMA CRUZI ANTIGEN, GENE ENCODING THEREFOR AND METHODS OF
; FILE REFERENCE: WPB 36400B
; CURRENT FILING DATE: US/09/138,735
; PRIOR APPLICATION NUMBER: 1998-08-24
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: FR 94/10132
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: probe/primer
US-09-138-735-8

Query Match 1.8%; Score 18; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 actctgcgcgtcgtcga 228
|||||
Db 18 ACTTCTGCCCTCTGTCA 1

RESULT 8
US-09-866-555-4980/c
; Sequence 4980, Application US/09866555
; GENERAL INFORMATION:
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Algate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Klee, Jennifer
; TITLE OF INVENTION: LUNG TUMOR-ASSOCIATED SEQUENCES AND METHODS OF USE
; FILE REFERENCE: 210121, 580
; CURRENT APPLICATION NUMBER: US/09/866,555
; CURRENT FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 20487
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4980
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature

LOCATION: (1)...(261)
OTHER INFORMATION: n = A,T,C or G
US-09-866-555-4980

Query Match 1.8%; Score 18; DB 6; Length 261;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 44 gtacgtcgctcgcgcgcg 61
|||||
Db 18 GTACTCTGCTCTGCTG 1

RESULT 9
US-09-863-524-3874
; Sequence 3874, Application US/09863524
; GENERAL INFORMATION:
; APPLICANT: Ito, Laura Y.
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN ROOTS
; FILE REFERENCE: PL-0003 US
; CURRENT APPLICATION NUMBER: US/09/863,524
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: 60/077,831
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 6814
; SOFTWARE: PERL Program
; SEQ ID NO 3874
; LENGTH: 297
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 700047828H1
US-09-863-524-3874

Query Match 1.8%; Score 18; DB 7; Length 297;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 386 cgccgcgcacgcacgcaggt 403
|||||
Db 247 cgccgcgcacgcacgcaggt 264

RESULT 10
US-09-764-891-1027/c
; Sequence 1027, Application US/09764891
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1027
; LENGTH: 328
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-1027

Query Match 1.8%; Score 18; DB 6; Length 328;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 actctgcgcgtcgtcga 228
|||||

Db 251 ACTTCTGCCGCTGCTGCA 234

RESULT 11

US-09-540-213-5492
; Sequence 5492, Application US/09540213
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF LIVER AND PANCREAS
; FILE REFERENCE: PD-1031 CIP
; CURRENT APPLICATION NUMBER: US/09/540,213
; CURRENT FILING DATE: 2000-03-31
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 53844
; SEQ ID NO 5492
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00532370
; NAME/KEY: unsure
; LOCATION: 4, 8, 304, 307
; OTHER INFORMATION: a, t, c, g, or other
US-09-540-213-5492

Query Match

1.8%; Score 18; DB 7; Length 348;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 cccctgcacccccctg 247
|||||

Db 29 cccctgcacccccctg 46

RESULT 12

US-09-667-188A-4578/c
; Sequence 4578, Application US/09667188A
; GENERAL INFORMATION:
; APPLICANT: Andersen, Scott E.
; APPLICANT: Masucci, James D.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51464)B
; CURRENT APPLICATION NUMBER: US/09/667,188A
; CURRENT FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/155,006
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 8135
; SEQ ID NO 4578
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3281-014-P1-K1-E7
US-09-667-188A-4578

Query Match

1.8%; Score 18; DB 6; Length 371;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 274 cctctgcgcgcgcgcga 291
|||||

Db 153 cctctgcgcgcgcgcga 136

RESULT 13

US-09-804-730-6853/c
; Sequence 6853, Application US/09804730
; GENERAL INFORMATION:
; APPLICANT: Andersen, Scott E.
; APPLICANT: Hammond-Kosack, Kim
; APPLICANT: Masucci, James D.
; APPLICANT: Urban, Martin
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51838)B
; CURRENT APPLICATION NUMBER: US/09/804,730
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: US 60/189,657
; PRIOR FILING DATE: 2000-03-15
; NUMBER OF SEQ ID NOS: 22828
; SEQ ID NO 6853
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Triticum aestivum
; OTHER INFORMATION: Clone ID: LIB3351-055-P1-K1-F5
US-09-804-730-6853

Query Match

1.8%; Score 18; DB 6; Length 378;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 cacacgcctctcccg 633
|||||

Db 173 CACACGCCTCTCCCG 156

RESULT 14

US-60-255-619-14255/c
; Sequence 14255, Application US/60255619
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)A
; CURRENT APPLICATION NUMBER: US/60/255,619
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 14255
; LENGTH: 380
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; OTHER INFORMATION: Clone ID: LIB3829-031-06-N6-D5
US-60-255-619-14255

Query Match

1.8%; Score 18; DB 8; Length 380;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 641 caacgatgcgttttg 658
|||||

Db 326 CAACGATGCGCTTTTG 309

RESULT 15

US-09-864-761-52/c
; Sequence 52, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

? GENERAL INFORMATION:
? APPLICANT: Hyseq, Inc.
? TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
? TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
? FILE REFERENCE: 20411-757
? CURRENT APPLICATION NUMBER: US/09/904, 809
? CURRENT FILING DATE: 2001-07-12
? PRIOR APPLICATION NUMBER: 09/234, 611
? PRIOR FILING DATE: 1999-01-22
? NUMBER OF SEQ ID NOS: 21025
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 19378
? LENGTH: 476
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)...(476)
? OTHER INFORMATION: n = A,T,C or G
US-09-904-809-19378

```

```
Query Match      1.8%; Score 18; DB 6; Length 476;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

| Qy | 44 | gragclgagctcfcgctg | 61 |
|----|-----|--------------------|-----|
| | | | |
| Db | 384 | gragctgcgtccrfgctg | 367 |

```

RESULT 17
US-09-909-627-11623/C
: Sequence 11623, Application US/09909627
: GENERAL INFORMATION:
: APPLICANT: hyseq, Inc.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
: TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
: FILE REFERENCE: 20411-766
: CURRENT APPLICATION NUMBER: US/09/909,627
: CURRENT FILING DATE: 2001-07-19
: PRIOR APPLICATION NUMBER: 09/277,127
: PRIOR FILING DATE: 1999-03-23
: NUMBER OF SEQ ID NOS: 23680
: SOFTWARE: fastseq for Windows Version 3.0
: SEQ ID NO 11623
: LENGTH: 476
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(476)
: OTHER INFORMATION: n = A,T,C or G
US-09-909-627-11623

```

Query Match . 1.8%; Score 18; DB 6; Length 476;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| Qy | 44 | gtagctgctcctgctg | 61 |
|----|-----|------------------|-----|
| Db | 384 | gtagctgctgctgctg | 367 |

RESULT 18
US-09-864-761-16893/c
; Sequence 16893, Application US/09864761/c
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 16893
LENGTH: 478
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC005921.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6.9
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
OTHER INFORMATION: SWISSPROT HIT: P50555, EVALUATE 5.60e-01
OTHER INFORMATION: EST HUMAN HIT: BE741337.1, EVALUATE 0.00e+00
OTHER INFORMATION: NT HIT: X99000.1, EVALUATE 3.00e-01
US-09-864-761-16893

Query Match 1.8%; Score 18; DB 7; Length 478;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 cccctgcacacccctg 247
|||||
DB 298 CCCCTGCATCACCCCTG 281

RESULT 19
US-09-840-145-13709/C
Sequence 13709, Application US/09840145
GENERAL INFORMATION:
APPLICANT: Adams, et. al.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products
FILE REFERENCE: P014C1
CURRENT APPLICATION NUMBER: US/09/840,145
PRIOR FILING DATE: 2001-04-24
PRIOR APPLICATION NUMBER: 08/276,163
PRIOR FILING DATE: 1994-07-15
NUMBER OF SEQ ID NOS: 15314
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13709
LENGTH: 506
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (11)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (30)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (44)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (46)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (51)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (60)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (66)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (148)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (185)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (306)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (321)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (345)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (359)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (364)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (375)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (385)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (425)
OTHER INFORMATION: n equals a,t,g, or c

```
; NAME/KEY: misc_feature
; LOCATION: (427)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (428)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (430)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (440)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (443)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (470)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (473)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (476)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (478)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (489)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (492)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (494)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (501)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-840-145-13709

Query Match          1.8%; Score 18; DB 7; Length 506;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 acctctgcgcgtgcga 228
    |||
Db 243 ACTTCTGCCGCTGCTGCA 226

RESULT 20
US-60-312-544-4166/C
; Sequence 4166, Application US/60312544
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)A
; CURRENT APPLICATION NUMBER: US/60/312,544
; CURRENT FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 10730
; SEQ ID NO 4166
; LENGTH: 1431
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(778)
```

```
; OTHER INFORMATION: Clone ID: LIB3205-113-E12.FLI
US-60-312-544-4166

Query Match          1.8%; Score 18; DB 8; Length 1431;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 386 cgcgcgcacacagagt 403
    |||
Db 542 CGCCGCGCACAGAGGT 525

RESULT 21
PCT-US01-16450-789/C
; Sequence 789, Application PC/FUS0116450
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P413PCT
; CURRENT APPLICATION NUMBER: PCT/US01/16450
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2820
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 789
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (158)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (749)
; OTHER INFORMATION: n equals a,t,g, or c
; PCT-US01-16450-789

Query Match          1.8%; Score 18; DB 1; Length 1485;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 gtacgtgcctcgtcgtc 61
    |||
Db 669 GTAGCTGCGTCGCTGCTG 652

RESULT 22
US-09-867-716-17510
; Sequence 17510, Application US/09867716
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Polynucleotides, Polypeptides Expressed
; FILE REFERENCE: 11000.1009c1
; CURRENT APPLICATION NUMBER: US/09/867,716
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 09/215,179
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 19860
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17510
; LENGTH: 1595
; TYPE: DNA
; ORGANISM: Pinus radiata
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1595)
; OTHER INFORMATION: n = A,T,C or G
; US-09-867-716-17510
```

Query Match 1.8%: Score 18; DB 7; Length 1595;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 814 acagagctgcagcagc 831
|||||
Db 1054 acagagctgcagcagc 1071

RESULT 23
US-09-836-472-1298/c
; Sequence 1298, Application US/09836472
; GENERAL INFORMATION:
; APPLICANT: Geating, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; FILE REFERENCE: 1600.1033-002
; CURRENT APPLICATION NUMBER: US/09/836,472
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 60/090,245
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 09/338,663
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 1364
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1298
; LENGTH: 2817
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2817)
; OTHER INFORMATION: n = A,T,C or G
US-09-836-472-1298

Query Match 1.8%: Score 18; DB 6; Length 2817;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 gtacgtcgctcgtcgtg 61
|||||
Db 1980 gtacgtcgctcgtcgtg 1963

RESULT 24
US-09-764-905-28829/c
; Sequence 28829, Application US/09764905
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC004
; CURRENT APPLICATION NUMBER: US/09/764,905
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290

;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/225,757
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/226,868
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/216,647
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/225,267
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/216,880
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/225,270
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/251,869
;; PRIOR FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/235,834
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/234,274
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: 60/234,223
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: 60/228,924
;; PRIOR FILING DATE: 2000-08-30
;; PRIOR APPLICATION NUMBER: 60/224,518
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/236,369
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/224,519
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/220,964
;; PRIOR FILING DATE: 2000-07-26
;; PRIOR APPLICATION NUMBER: 60/241,809
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/249,299
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/236,327
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/241,785
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/244,617
;; PRIOR FILING DATE: 2000-11-01
;; PRIOR APPLICATION NUMBER: 60/225,268
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/236,368
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/251,856
;; PRIOR FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/251,868
;; PRIOR FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/229,344
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/234,997
;; PRIOR FILING DATE: 2000-09-25
;; PRIOR APPLICATION NUMBER: 60/229,343
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,345
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,287
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,513
;; PRIOR FILING DATE: 2000-09-05
;; PRIOR APPLICATION NUMBER: 60/231,413
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/229,509
;; PRIOR FILING DATE: 2000-09-05
;; PRIOR APPLICATION NUMBER: 60/236,367
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/237,039
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,038
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/236,370
;; PRIOR FILING DATE: 2000-09-29

PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241,787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,245
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,244
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,217
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,211
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,215
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,264
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,214
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,297
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232,400
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231,242
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,081
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080

PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,414
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,244
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,064
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/233,063
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,397
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,399
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,401
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/241,808
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,826
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,786
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,221
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,475
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/231,243
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,065
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,398

Query Match 1.8%; Score 18; DB 7; Length 15270;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 catccctgcatcacc 244
DB 11976 CATCCCTGCATCACC 11959

RESULT 25
US-09-764-905-40566/c
Sequence 40566, Application US/09764905
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC004
CURRENT FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/225,757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07

;; PRIOR APPLICATION NUMBER: 60/225,267
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/216,880
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/225,270
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/251,869
;; PRIOR FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/235,834
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/234,274
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: 60/234,223
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: 60/228,924
;; PRIOR FILING DATE: 2000-08-30
;; PRIOR APPLICATION NUMBER: 60/224,518
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/236,369
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/224,519
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/220,964
;; PRIOR FILING DATE: 2000-07-26
;; PRIOR APPLICATION NUMBER: 60/241,809
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/249,299
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/236,327
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/241,785
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/244,617
;; PRIOR FILING DATE: 2000-11-01
;; PRIOR APPLICATION NUMBER: 60/225,268
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/236,368
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/251,856
;; PRIOR FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/251,868
;; PRIOR FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: 60/229,344
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/234,997
;; PRIOR FILING DATE: 2000-09-25
;; PRIOR APPLICATION NUMBER: 60/229,343
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,345
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,287
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,513
;; PRIOR FILING DATE: 2000-09-05
;; PRIOR APPLICATION NUMBER: 60/231,413
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/229,509
;; PRIOR FILING DATE: 2000-09-05
;; PRIOR APPLICATION NUMBER: 60/236,367
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/237,039
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,038
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/236,370
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/236,802
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,037
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,040
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/240,960

;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/239,935
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: 60/239,937
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: 60/241,787
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/246,474
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/246,532
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/249,216
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,210
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/226,681
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/225,759
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/225,213
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/227,182
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/225,214
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/235,836
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/230,438
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/215,135
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: 60/225,266
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/249,218
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,208
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,213
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,212
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,207
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,245
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,244
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,217
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,211
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,215
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,264
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,214
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,297
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/232,400
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/231,242
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,081
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,080
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,414
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,244
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/233,064
;; PRIOR FILING DATE: 2000-09-14

```
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,065
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,398
```

```
Query Match      1.8%; Score 18; DB 7; Length 15270;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 227 catccctcatcacc 244
|||||
```

```
Db 11976 CATCCCTCATCACC 11959
```

```
RESULT 26
; Sequence 1271, Application US/09803736
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Levin, Irena M.
; APPLICANT: Norris, Susan R.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)D
; CURRENT APPLICATION NUMBER: US/09/803,736
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/534,859
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: Identified by Attorney Docket number 04983.0206CPUS01 38-10
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 1582
; SEQ ID NO 1271
; LENGTH: 62271
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-803-736-1271
```

```
Query Match      1.8%; Score 18; DB 6; Length 62271;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 779 aacaactctgaatc 796
|||||
```

```
Db 57285 aacaactctgaatc 57302
```

```
RESULT 27
US-09-803-736-896/c
; Sequence 896, Application US/09803736
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Levin, Irena M.
```

```
; APPLICANT: Norris, Susan R.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)D
; CURRENT APPLICATION NUMBER: US/09/803,736
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/534,859
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: Identified by Attorney Docket number 04983.0206CPUS01 38
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 1582
; SEQ ID NO 896
; LENGTH: 78596
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-803-736-896
```

```
Query Match      1.8%; Score 18; DB 6; Length 78596;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 930 caacatcaggaagggt 947
|||||
```

```
Db 41282 CAACATCAGGAGG 41265
```

```
RESULT 28
US-09-803-736-1270/c
; Sequence 1270, Application US/09803736
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Levin, Irena M.
; APPLICANT: Norris, Susan R.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)D
; CURRENT APPLICATION NUMBER: US/09/803,736
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/534,859
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: Identified by Attorney Docket number 04983.0206CPUS01 38
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 1582
; SEQ ID NO 1270
; LENGTH: 87045
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-803-736-1270
```

```
Query Match      1.8%; Score 18; DB 6; Length 87045;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 779 aacaactctgaatc 796
|||||
```

```
Db 67036 AACACTCTGATCTG 67019
```

```
RESULT 29
US-09-803-736-19
; Sequence 19, Application US/09803736
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Levin, Irena M.
; APPLICANT: Norris, Susan R.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)D
```



```
; CURRENT APPLICATION NUMBER: US/09/803,736
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/534,859
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: Identified by Attorney Docket number 04983.0206CPUS01 38-10
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 1582
; SEQ ID NO 19
; LENGTH: 91470
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-803-736-19
```

```
Query Match          1.8%; Score 18; DB 6; Length 91470;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 797 gtttagagcacagatga 814
|||||
Db 63394 gtttagagcacagatga 63411
```

```
RESULT 30
US-09-803-736-1397/c
; Sequence 1397, Application US/09803736
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Levin, Irena M.
; APPLICANT: Norris, Susan R.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Miegand, Roger C.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)D
; CURRENT APPLICATION NUMBER: US/09/803,736
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/534,859
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: Identified by Attorney Docket number 04983.0206CPUS01 38-10
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 1382
; SEQ ID NO 1397
; LENGTH: 96424
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-803-736-1397
```

```
Query Match          1.8%; Score 18; DB 6; Length 96424;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 202 accaatcgactctgcc 219
|||||
Db 81927 ACCAATCGACTCTGCC 81910
```

```
RESULT 31
US-09-803-736-1100
; Sequence 1100, Application US/09803736
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Levin, Irena M.
; APPLICANT: Norris, Susan R.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Miegand, Roger C.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)D
; CURRENT APPLICATION NUMBER: US/09/803,736
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/534,859
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: Identified by Attorney Docket number 04983.0206CPUS01 38-10
```

```
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 1582
; SEQ ID NO 1100
; LENGTH: 110684
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-803-736-1100
```

```
Query Match          1.8%; Score 18; DB 6; Length 110684;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 652 gtttgcgcgcagcagta 669
|||||
Db 34775 gtttgcgcgcagcagta 34792
```

```
RESULT 32
US-60-253-456-35126/c
; Sequence 35126, Application US/60253456
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka J
; TITLE OF INVENTION: Polynucleotides, isolated from
; TITLE OF INVENTION: plants, and methods for their use.
; FILE REFERENCE: 1054p1
; CURRENT APPLICATION NUMBER: US/60/253,456
; CURRENT FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 37096
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35126
; LENGTH: 88
; TYPE: DNA
; ORGANISM: Pinus radiata
US-60-253-456-35126
```

```
Query Match          1.7%; Score 17; DB 8; Length 88;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 842 cgcacaacacagccgaga 858
|||||
Db 74 CGCAAAACACAGCCGAGA 58
```

```
RESULT 33
US-60-253-457-47464/c
; Sequence 47464, Application US/60253457
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka J
; TITLE OF INVENTION: Polynucleotides, isolated from plants
; TITLE OF INVENTION: and methods for their use.
; FILE REFERENCE: 1054p2
; CURRENT APPLICATION NUMBER: US/60/253,457
; CURRENT FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 48893
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47464
; LENGTH: 119
; TYPE: DNA
; ORGANISM: Pinus radiata
US-60-253-457-47464
```

```
Query Match          1.7%; Score 17; DB 8; Length 119;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 842 cgcacaacacagccgaga 858
|||||
Db 100 CGCAAAACACAGCCGAGA 84
```


OTHER INFORMATION: Incyte ID No: rat00171041
US-09-540-213-52685

Query Match 1.7%; Score 17; DB 7; Length 241;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 aacagcagcagcgccaa 129
|||||
DB 71 aacagcagcagcgccaa 87

RESULT 39
US-60-253-457-26897/C
Sequence 26897, Application US/60253457
GENERAL INFORMATION:
APPLICANT: Havukala, Ilkka J
TITLE OF INVENTION: Polynucleotides, isolated from plants
TITLE OF INVENTION: and methods for their use.
FILE REFERENCE: 1054p2
CURRENT APPLICATION NUMBER: US/60/253,457
CURRENT FILING DATE: 2000-11-27
NUMBER OF SEQ ID NOS: 48893
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 26897
LENGTH: 250
TYPE: DNA
ORGANISM: pinus radiata
US-60-253-457-26897

Query Match 1.7%; Score 17; DB 8; Length 250;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 842 cgcacacacagcgaga 858
|||||
DB 91 CGCAACACACAGCCGAGA 75

RESULT 40
US-09-864-761-27318
Sequence 27318, Application US/09864761
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 27318
LENGTH: 262
TYPE: DNA
ORGANISM: Homo sapiens

OTHER INFORMATION: MAP TO AC004707.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.78
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.82
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.84
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.97
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.73
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.88
OTHER INFORMATION: NT HIT: g111434402, EVALUATE 1.00e-91
OTHER INFORMATION: EST_HUMAN HIT: BE891148.1, EVALUATE 2.00e-91
OTHER INFORMATION: SWISSPROT HIT: P16815, EVALUATE 4.20e+00
US-09-864-761-27318

Query Match 1.7%; Score 17; DB 7; Length 262;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 812 tgacagagctgcagcag 828
|||||
DB 127 tgacagagctgcagcag 143

RESULT 41
US-09-540-213-44587
Sequence 44587, Application US/09540213
GENERAL INFORMATION:
APPLICANT: Seilhamer, Jeffrey J.
APPLICANT: Deleageane, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullaby, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES OF LIVER AND PANCREAS
FILE REFERENCE: PD-1031 CIP
CURRENT APPLICATION NUMBER: US/09/540,213
CURRENT FILING DATE: 2000-03-31
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 53844
SEQ ID NO 44587
LENGTH: 264
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No: rat00163047
US-09-540-213-44587

Query Match 1.7%; Score 17; DB 7; Length 264;

Best Local Similarity 100.0%; Pred. No. 93;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 113 aacagcagccagcgccaa 129
|||||
Db 71 aacagcagccagcgccaa 87

RESULT 42
US-60-253-456-17545/C
; Sequence 17545, Application US/60253456
; GENERAL INFORMATION:
; APPLICANT: Havukkaia, Ilkka J
; TITLE OF INVENTION: Polynucleotides, isolated from
; FILE REFERENCE: 1054p1
; CURRENT APPLICATION NUMBER: US/60/253,456
; CURRENT FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 37096
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 17545
; LENGTH: 268
; TYPE: DNA
; ORGANISM: Pinus radiata
US-60-253-456-17545

Query Match 1.7%; Score 17; DB 8; Length 268;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 842 cgcacaacagcgcgaga 858
|||||
Db 122 CCACAACACAGCCGAGA 106

RESULT 43
US-09-912-292-45542
; Sequence 45542, Application US/09912292
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 101
; FILE REFERENCE: PO-101
; CURRENT APPLICATION NUMBER: US/09/912,292
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 08/103,746
; PRIOR FILING DATE: 1993-08-09
; PRIOR APPLICATION NUMBER: 09/859,417
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 08/103,743
; PRIOR FILING DATE: 1993-08-09
; PRIOR APPLICATION NUMBER: 09/741,827
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/813,154
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 08/196,482
; PRIOR FILING DATE: 1994-02-15
; PRIOR APPLICATION NUMBER: 09/783,587
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 08/196,481
; PRIOR FILING DATE: 1994-02-15
; PRIOR APPLICATION NUMBER: 08/220,662
; PRIOR FILING DATE: 1994-03-31
; PRIOR APPLICATION NUMBER: 08/220,661
; PRIOR FILING DATE: 1994-03-31
; PRIOR APPLICATION NUMBER: 08/275,627
; PRIOR FILING DATE: 1994-07-15
; PRIOR APPLICATION NUMBER: 08/345,704
; PRIOR FILING DATE: 1994-11-21
; PRIOR APPLICATION NUMBER: 09/859,662
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 08/401,881
; PRIOR FILING DATE: 1995-03-10

; PRIOR APPLICATION NUMBER: 09/782,161
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 08/790,776
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: 60/010,803
; PRIOR FILING DATE: 1996-01-30
; PRIOR APPLICATION NUMBER: 08/799,180
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: 60/011,618
; PRIOR FILING DATE: 1996-02-13
; PRIOR APPLICATION NUMBER: 08/803,609
; PRIOR FILING DATE: 1997-02-21
; PRIOR APPLICATION NUMBER: 60/011,985
; PRIOR FILING DATE: 1996-02-21
; PRIOR APPLICATION NUMBER: 08/971,050
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: 60/031,038
; PRIOR FILING DATE: 1996-11-18
; PRIOR APPLICATION NUMBER: 08/975,985
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/031,221
; PRIOR FILING DATE: 1996-11-25
; PRIOR APPLICATION NUMBER: 08/985,366
; PRIOR FILING DATE: 1997-12-04
; PRIOR APPLICATION NUMBER: 09/842,827
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/032,782
; PRIOR FILING DATE: 1996-12-06
; PRIOR APPLICATION NUMBER: 09/076,898
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/047,004
; PRIOR FILING DATE: 1997-05-13
; PRIOR APPLICATION NUMBER: 09/078,614
; PRIOR FILING DATE: 1998-05-14
; PRIOR APPLICATION NUMBER: 60/046,498
; PRIOR FILING DATE: 1997-05-14
; PRIOR APPLICATION NUMBER: 09/189,834
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: 60/065,454
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 09/366,690
; PRIOR FILING DATE: 1999-08-04
; PRIOR APPLICATION NUMBER: 60/095,484
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 09/342,216
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/091,171
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: 09/487,572
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/116,678
; PRIOR FILING DATE: 1999-01-21
; PRIOR APPLICATION NUMBER: 09/497,772
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/118,784
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 09/504,577
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/120,434
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 09/517,011
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/122,447
; PRIOR FILING DATE: 1999-03-02
; PRIOR APPLICATION NUMBER: 09/628,857
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/145,956
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 09/628,858

```

; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/145,955
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 09/741,033
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 09/773,518
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/171,621
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/796,765
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 09/867,682
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/187,016
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/225,269
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,629
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 55553
; SEQ ID NO 45542
; LENGTH: 273
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28)..(28)
; OTHER INFORMATION: n is equal to a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (43)..(43)
; OTHER INFORMATION: n is equal to a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (52)..(53)
; OTHER INFORMATION: n is equal to a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (59)..(59)
; OTHER INFORMATION: n is equal to a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (105)..(105)
; OTHER INFORMATION: n is equal to a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (122)..(123)
; OTHER INFORMATION: n is equal to a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (150)..(151)
; OTHER INFORMATION: n is equal to a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (153)..(154)
; OTHER INFORMATION: n is equal to a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (157)..(158)
; OTHER INFORMATION: n is equal to a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (170)..(171)
; OTHER INFORMATION: n is equal to a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (230)..(230)
; OTHER INFORMATION: n is equal to a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (257)..(257)
; OTHER INFORMATION: n is equal to a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (261)..(261)
; OTHER INFORMATION: n is equal to a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (264)..(264)
; OTHER INFORMATION: n is equal to a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (273)..(273)
; OTHER INFORMATION: n is equal to a,t,g, or c
; US-09-912-292-45542

```

```

Query Match      1.7%; Score 17; DB 7; Length 273;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 812 tgcacagctgcagcag 828
|||||
Db 10 tgcacagctgcagcag 26

```

```

RESULT 44
US-60-253-378-16751/c
; Sequence 16751, Application US/60253378
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka J
; TITLE OF INVENTION: Polynucleotides isolated from plants
; TITLE OF INVENTION: and methods for their use.
; FILE REFERENCE: 1054p3
; CURRENT APPLICATION NUMBER: US/60/253,378
; CURRENT FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 40367
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16751
; LENGTH: 276
; TYPE: DNA
; ORGANISM: Pinus radiata
US-60-253-378-16751

```

```

Query Match      1.7%; Score 17; DB 8; Length 276;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 842 cgcacaacagccgaga 858
|||||
Db 104 CGCAACACAGCCGAGA 88

```

```

RESULT 45
US-60-253-457-27422/c
; Sequence 27422, Application US/60253457
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka J
; TITLE OF INVENTION: Polynucleotides, isolated from plants
; TITLE OF INVENTION: and methods for their use.
; FILE REFERENCE: 1054p2
; CURRENT APPLICATION NUMBER: US/60/253,457
; CURRENT FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 48893
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27422
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Pinus radiata
US-60-253-457-27422

```

```

Query Match      1.7%; Score 17; DB 8; Length 278;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 842 cgcacaacagccgaga 858
|||||
Db 115 CGCAACACAGCCGAGA 99

```

```

RESULT 46
US-60-253-457-35043/c
; Sequence 35043, Application US/60253457
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka J
; TITLE OF INVENTION: Polynucleotides, isolated from plants
; TITLE OF INVENTION: and methods for their use.
; FILE REFERENCE: 1054p2

```

```

: CURRENT APPLICATION NUMBER: US/60/253,457
: CURRENT FILING DATE: 2000-11-27
: NUMBER OF SEQ ID NOS: 48893
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 35043
: LENGTH: 279
: TYPE: DNA
: ORGANISM: Pinus radiata
US-60-253-457-35043
```

```

Query Match          1.7%: Score 17; DB 8; Length 279;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 842 cgcacacacagccgaga 858
    |||||||
Db 122 CGCAACACACGCCGAGA 106
```

```

RESULT 47
US-09-724-750-1826/C
: Sequence 1826, Application US/09724750
: GENERAL INFORMATION:
: APPLICANT: Havukkala, Ilkka
: TITLE OF INVENTION: Polynucleotide, Materials Incorporating
: FILE REFERENCE: 11000.1049AU
: CURRENT APPLICATION NUMBER: US/09/724,750
: CURRENT FILING DATE: 2000-11-28
: PRIOR APPLICATION NUMBER: 60/171,431
: PRIOR FILING DATE: 1999-12-21
: NUMBER OF SEQ ID NOS: 25199
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1826
: LENGTH: 280
: TYPE: DNA
: ORGANISM: Pinus radiata
US-09-724-750-1826
```

```

Query Match          1.7%: Score 17; DB 7; Length 280;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 842 cgcacacacagccgaga 858
    |||||||
Db 130 CGCAACACACGCCGAGA 114
```

```

RESULT 48
US-09-863-524-1238
: Sequence 1238, Application US/09863524
: GENERAL INFORMATION:
: APPLICANT: Ito, Laura Y.
: APPLICANT: Gooding, Douglas H.
: APPLICANT: Sherman, Bradley K.
: TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN ROOTS
: FILE REFERENCE: PL-0003 US
: CURRENT APPLICATION NUMBER: US/09/863,524
: CURRENT FILING DATE: 2001-05-22
: PRIOR APPLICATION NUMBER: 60/077,831
: PRIOR FILING DATE: 1998-03-13
: NUMBER OF SEQ ID NOS: 6814
: SOFTWARE: PERL Program
: SEQ ID NO 1238
: LENGTH: 289
: TYPE: DNA
: ORGANISM: Zea mays
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 700027505H1
NAME/KEY: unsure
```

```

: LOCATION: 118-146, 183
: OTHER INFORMATION: a, t, c, g, or other
US-09-863-524-1238
```

```

Query Match          1.7%: Score 17; DB 7; Length 289;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 698 cgcctgagcgcgcctgag 714
    |||||||
Db 50 cgcctgagcgcgcctgag 66
```

```

RESULT 49
US-60-253-457-35140/C
: Sequence 35140, Application US/60253457
: GENERAL INFORMATION:
: APPLICANT: Havukkala, Ilkka J
: TITLE OF INVENTION: Polynucleotides, Isolated from plants
: FILE REFERENCE: 1054P2
: CURRENT APPLICATION NUMBER: US/60/253,457
: CURRENT FILING DATE: 2000-11-27
: NUMBER OF SEQ ID NOS: 48893
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 35140
: LENGTH: 291
: TYPE: DNA
: ORGANISM: Pinus radiata
US-60-253-457-35140
```

```

Query Match          1.7%: Score 17; DB 8; Length 291;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 842 cgcacacacagccgaga 858
    |||||||
Db 122 CGCAACACACGCCGAGA 106
```

```

RESULT 50
US-09-540-213-33702
: Sequence 33702, Application US/09540213
: GENERAL INFORMATION:
: APPLICANT: Selbame, Jeffrey J.
: APPLICANT: Deleane, Angelo M.
: APPLICANT: Stuart, Susan G.
: APPLICANT: Stuve, Laura L.
: APPLICANT: Mullahy, Sara J.
: APPLICANT: Naughton, Rebecca E.
: TITLE OF INVENTION: POLYNUCLEOTIDES OF LIVER AND PANCREAS
: FILE REFERENCE: PD-1031 CIP
: CURRENT APPLICATION NUMBER: US/09/540,213
: CURRENT FILING DATE: 2000-03-31
: Prior application data removed - consult PALM or file wrapper
: NUMBER OF SEQ ID NOS: 53844
: SEQ ID NO 33702
: LENGTH: 300
: TYPE: DNA
: ORGANISM: Rattus norvegicus
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: rat00076798
US-09-540-213-33702
```

```

Query Match          1.7%: Score 17; DB 7; Length 300;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 113 aacacagccagccgcaa 129
```

Mon Sep 24 09:36:18 2001

us-09-138-735-1_copy_1232_2207.oli5.rnpn

Page 25

Db 76 aacagcagccagccaa 92

Search completed: September 21, 2001, 13:39:08
Job time: 909 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 00:36:20 : Search time 3479.37 Seconds
(without alignments)
93.357 Million cell updates/sec

Title: US-09-138-735-12
Perfect score: 21
Sequence: 1 cgaagagaccatgaacaactt 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba1:*
- 2: gb_ba2:*
- 3: gb_ba3:*
- 4: gb_in1:*
- 5: gb_in2:*
- 6: gb_in3:*
- 7: gb_om:*
- 8: gb_ov:*
- 9: gb_pat1:*
- 10: gb_pat2:*
- 11: gb_ph:*
- 12: gb_p11:*
- 13: gb_p12:*
- 14: gb_p13:*
- 15: gb_p14:*
- 16: gb_ba1:*
- 17: gb_ba2:*
- 18: gb_fun:*
- 19: gb_htg_hum:*
- 20: gb_htgo_inv:*
- 21: gb_htgo_inv:*
- 22: gb_htg_hum1:*
- 23: gb_htg_hum2:*
- 24: gb_htg_hum3:*
- 25: gb_htg_hum4:*
- 26: gb_htg_hum5:*
- 27: gb_htg_hum6:*
- 28: gb_htg_hum7:*
- 29: gb_htg_hum8:*
- 30: gb_htg_inv1:*
- 31: gb_htg_inv2:*
- 32: gb_htg_other:*
- 33: gb_htg_rod:*
- 34: gb_hum1:*
- 35: gb_hum2:*
- 36: gb_hum3:*
- 37: gb_hum4:*
- 38: gb_hum5:*
- 39: gb_hum6:*
- 40: gb_hum7:*
- 41: gb_in:*
- 42: gb_om:*
- 43: gb_or:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| 1 | 21 | 100.0 | 21 | A48921 | A48921 Sequence 12 |
| 2 | 21 | 100.0 | 21 | AR047930 | AR047930 Sequence 1 |
| 3 | 21 | 100.0 | 3402 | A48910 | A48910 Sequence 1 |
| 4 | 21 | 100.0 | 3402 | AR047920 | AR047920 Sequence 1 |
| 5 | 21 | 100.0 | 3402 | TCU24190 | U24190 Trypanosoma |
| 6 | 19 | 90.5 | 67442 | AC015691 | AC015691 Homo sapi |
| 7 | 18.4 | 87.6 | 64744 | AC090694 | AC090694 Homo sapi |
| 8 | 18.4 | 87.6 | 215798 | AL161719 | AL161719 Homo sapi |

| | | | | | | | |
|---|----|------|------|---------|----|------------|---------------------|
| C | 9 | 18 | 85.7 | 147042 | 62 | AC0116647 | AC0116647 Homo sapi |
| C | 10 | 18 | 85.7 | 1999958 | 69 | AC0260683 | AC0260683 Homo sapi |
| C | 11 | 17.8 | 84.8 | 2984 | 3 | PST012352 | AT012352 Pseudomon |
| C | 12 | 17.8 | 84.8 | 6052 | 1 | AF121970 | AF121970 Pseudomon |
| C | 13 | 17.8 | 84.8 | 6921 | 1 | AF027768 | AF027768 Serratia |
| C | 14 | 17.8 | 84.8 | 174376 | 64 | AC0165627 | AC0165627 Homo sapi |
| C | 15 | 17.8 | 84.8 | 205638 | 71 | AC0406927 | AC0406927 Homo sapi |
| C | 16 | 17.8 | 84.8 | 228434 | 62 | AC0122205 | AC0122205 Homo sapi |
| C | 17 | 17.4 | 82.9 | 313 | 54 | HS355TH1 | Z5152 H.sapiens (|
| C | 18 | 17.4 | 82.9 | 128235 | 93 | HSJ9955C14 | AL050340 Human DNA |
| C | 19 | 17.4 | 82.9 | 1359583 | 93 | HSJ5959K12 | AL0596829 Human DNA |
| C | 20 | 17.4 | 82.9 | 146431 | 75 | AC079098 | AC079098 Homo sapi |
| C | 21 | 17.4 | 82.9 | 148030 | 78 | AL137803 | AL137803 Homo sapi |
| C | 22 | 17.4 | 82.9 | 155356 | 73 | AC068308 | AC068308 Homo sapi |
| C | 23 | 17.4 | 82.9 | 170887 | 79 | AL161910 | AL161910 Homo sapi |
| C | 24 | 17.4 | 82.9 | 173381 | 71 | AC040918 | AC040918 Homo sapi |
| C | 25 | 17.4 | 82.9 | 196941 | 64 | AC071062 | AC071062 Homo sapi |
| C | 26 | 17.4 | 82.9 | 312420 | 82 | AL500233 | AL500233 Homo sapi |
| C | 27 | 17 | 81.0 | 165443 | 76 | AC079939 | AC079939 Mus muscu |
| C | 28 | 17 | 81.0 | 191352 | 74 | AC073565 | AC073565 Mus muscu |
| C | 29 | 16.8 | 80.0 | 1590 | 6 | CEU25632 | U25632 Caenorhabdi |
| C | 30 | 16.8 | 80.0 | 2141 | 89 | AK001946 | AK001946 Homo sapi |
| C | 31 | 16.8 | 80.0 | 2158 | 93 | HSM601650 | AL136682 Homo sapi |
| C | 32 | 16.8 | 80.0 | 2163 | 9 | AX056419 | AX056419 Sequence |
| C | 33 | 16.8 | 80.0 | 2163 | 89 | AF113249 | AF113249 Homo sapi |
| C | 34 | 16.8 | 80.0 | 2197 | 91 | BC001221 | BC001221 Homo sapi |
| C | 35 | 16.8 | 80.0 | 15033 | 63 | AC012888 | AC012888 Drosophill |
| C | 36 | 16.8 | 80.0 | 31800 | 6 | CET21B10 | Z68313 Caenorhabdi |
| C | 37 | 16.8 | 80.0 | 50550 | 11 | AF271693 | AF271693 mycobacte |
| C | 38 | 16.8 | 80.0 | 63409 | 67 | AC032982 | AC032982 Homo sapi |
| C | 39 | 16.8 | 80.0 | 65843 | 69 | AC025903 | AC025903 Homo sapi |
| C | 40 | 16.8 | 80.0 | 85560 | 65 | AC019835 | AC019835 Drosophill |
| C | 41 | 16.8 | 80.0 | 86719 | 91 | AP000885 | AP000885 Homo sapi |
| C | 42 | 16.8 | 80.0 | 94378 | 13 | ATF15B8 | AL049660 Arabidops |
| C | 43 | 16.8 | 80.0 | 100834 | 90 | AL160411 | AL160411 Human DNA |
| C | 44 | 16.8 | 80.0 | 104087 | 87 | AC008926 | AC008926 Homo sapi |
| C | 45 | 16.8 | 80.0 | 104549 | 62 | AC011384 | AC011384 Homo sapi |

| | | | |
|-----------------------------|--|------------------------|-------------------|
| OY | 1 | CGAAGAGACCATgaacactt | 21 |
| | | | |
| Dd | 1 | CGAAGAGACCATGAACAACCTT | 21 |
| RESULT | 2 | | |
| LOCUS | AR047930 | | |
| DEFINITION | AR047930 | 21 bp | DNA |
| ACCSSION | Sequence 12 from patent US 5620864. | | PAT |
| VERSION | AR047930 | | |
| KEYWORDS | AR047930.1 GI:5970273 | | |
| SOURCE | Unknown. | | |
| ORGANISM | Unknown. | | |
| REFERENCE | Unclassified. | | |
| AUTHORS | 1 (bases 1 to 21) | | |
| TITLE | Paranohs-Baccala,G., Iesenechal,M. and Jolivet,M. | | |
| JOURNAL | Trypanosoma cruzi antigen, gene encoding therefor and methods of | | |
| FEATURES | detecting and treating chagas disease | | |
| | Patent: US 5820864-A 12 13-OCT-1998; | | |
| | Location/Qualifiers | | |
| | 1..21 | | |
| BASE COUNT | 9 a | 5 c | 4 g |
| ORIGIN | 3 t | | |
| | /organism="unknown" | | |
| Query Match | 100.0%; | Score 21; | DB 9; Length 21; |
| Best Local Similarity | 100.0%; | Pred. NO. 2; | |
| Matches 21; Conservative 0; | Mismatches | 0; | Indels 0; Gaps 0; |
| OY | 1 | CGAAGAGACCAtaaacactt | 21 |
| | | | |
| Dd | 1 | CGAAGAGACCATGAACAACCTT | 21 |
| RESULT | 3 | | |

| ALIGNMENTS | | | | | | | | | |
|-----------------------|--|--------|--------------|-------|------------|----|------|----|--|
| RESULT | 1 | | | | | | | | |
| LOCUS | A48921 | | | | | | | | |
| DEFINITION | A48921 | 21 bp | DNA | | | | | | |
| ACCESSION | Sequence 12 from Patent WO9605312. | | | | | | | | |
| VERSION | A48921 | | | | | | | | |
| KEYWORDS | A48921.1 GI:2302580 | | | | | | | | |
| SOURCE | unidentified. | | | | | | | | |
| ORGANISM | unidentified | | | | | | | | |
| REFERENCE | 1 (bases 1 to 21) | | | | | | | | |
| AUTHORS | Paranhos-Paccala,G., Lesenechal,M. and Jolivet,M. | | | | | | | | |
| TITLE | NOVEL TRYPANOSOMA CRUZI ANTIGEN, AND GENE CODING THEREFOR; THEIR APPLICATION TO THE DETECTION OF CHAGAS' DISEASE | | | | | | | | |
| JOURNAL | Patent: WO 9605312-A 12 22-FEB-1996; | | | | | | | | |
| COMMENT | BIO MERIEUX (FR) | | | | | | | | |
| FEATURES | Other publication CA 2173957 960322 | | | | | | | | |
| SOURCE | Other publication AU 3169195 960307 | | | | | | | | |
| | Other publication FR 2723589 960216. | | | | | | | | |
| | 1. 21 | | | | | | | | |
| | location/Qualifiers | | | | | | | | |
| | /organism="unidentified" | | | | | | | | |
| | /db_xref="taxon:32644" | | | | | | | | |
| BASE COUNT | 9 a | 5 c | 4 g | 3 t | | | | | |
| ORIGIN | | | | | | | | | |
| Query Match | | 100.0% | Score 21; | DB 9; | Length 21; | | | | |
| Best Local Similarity | | 100.0% | Pred. No. 2; | | | | | | |
| Matches 21; | Conservative | 0; | Mismatches | 0; | Indels | 0; | Gaps | 0; | |

[illegible]

```

RESULT 4
LOCUS AR047920 3402 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5820864.
ACCESSION AR047920
VERSION AR047920.1 GI:5970263
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1.3402
/organism="unknown"
BASE COUNT 888 a 821 c 956 g 737 t
ORIGIN
Query Match 100.0%; Score 21; DB 9; Length 3402;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 cgaagagaccatgacaactt 21
|||||
Db 1997 CGAAGAGACCATGACCACTT 2017

RESULT 5
LOCUS TC024190 3402 bp mRNA INV 04-AUG-1997
DEFINITION Trypanosoma cruzi Tc40 antigen (Tc40) mRNA, complete cds.
ACCESSION U24190
VERSION U24190.1 GI:790645
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1.3402
/organism="Trypanosoma cruzi"
/db_xref="taxon:5693"
/dev_stage="epimastigote"
/misc_RNA
1.35
/gene="Tc40"
/note="spliced leader"
1.3402
/gene="Tc40"
/feature="Tc40"
266..3013
/gene="Tc40"
/codon_start=1
/product="Tc40 antigen"
/protein_id="AAC47657.1"

```

```

/db_xref="GI:790646"
/translation="MTVTVDLFFNNAKPSNNEGRVMSVDAATFNEPPEAORVLAADSOY
LAYMKRRHVLRVYKRSNLTKGTVAHSPKPTTHAVKFNVRNSNAVSGKGEFFWVVT
DETASNCKPDLARLTLYKYVFKIADPTTICSPFTINAEQRPDLILVETQAIIID
SSLIERDVSLRATLORCTTLRTLTLPVSENSLCSVSGGWNFTFTETRYACT
LRNSTPDMACCEGEPEVKALHLDAVYENVSVLVAATKGVOMLTGLVAEPMLRK
FVIGDSIVAMESSRETFVFDKROKALVNNHSPNFCVTHYMPCCQVRNGFCENRT
ADGSCVLAADQGNRLTIFHLRSRREDOOPGOKTSVATKPGCVSGSTNSHTNT
TSAAPASPSPPVSPAPAKAAPPAARSAEPHVGSKITIANLVNOLINVOARSVSTG
APATRTSTAVSTTTAPORTSPYCHNGRPVTAGLVANSGASAASSPTAAKPGCEK
ASACETSSVAINATRPALHNASLPDAPFDGVLAAAYOSFEVHOSLERLESTVINT
SRVLKLPDITLRDHEQLNLNGLAOMTELQOSRPTQOPORDPSAKSVFEETVL
LIADSLRNITKGVKGVNEAIMLHDEHVAHGNRLRQOKNIKSRDEALKEST
TOFTAOULTOVENVKRELAVLGSINGSLSVKENASLOKEINSIMSGVLDENMR
MBEELCTRESVAKRKATMPDSILHATSSPFGSRAPETTLATATLSMREOYRGLE
YMLAQQPSLLRLSLITRENNAYSELINVERPNWCVSYLIQLEAATYAEKE
VVGVADILILSERDQIAONGALGSLTTAMKAFERQARSSTSSFLQCLNKLRLIO
S"
BASE COUNT 889 a 818 c 958 g 737 t
ORIGIN
Query Match 100.0%; Score 21; DB 96; Length 3402;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 cgaagagaccatgacaactt 21
|||||
Db 1997 CGAAGAGACCATGACCACTT 2017

RESULT 6
LOCUS AC015691 67442 bp DNA HTG 17-NOV-1999
DEFINITION Homo sapiens chromosome 11 clone RP11-54f1 map 11, LOW-PASS
SEQUENCE SAMPLING.
ACCESSION AC015691
VERSION AC015691.1 GI:6446996
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1.3402
/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="clonal"
/misc_RNA
1.35
/gene="Tc40"
/note="spliced leader"
1.3402
/gene="Tc40"
/feature="Tc40"
266..3013
/gene="Tc40"
/codon_start=1
/product="Tc40 antigen"
/protein_id="AAC47657.1"

```

* NOTE: This record contains 78 individual

* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 867: contig of 867 bp in length
* gap of unknown length
* 868 1672: contig of 805 bp in length
* gap of unknown length
* 1673 2513: contig of 841 bp in length
* gap of unknown length
* 2514 3397: contig of 884 bp in length
* gap of unknown length
* 3398 4293: contig of 896 bp in length
* gap of unknown length
* 4294 5163: contig of 870 bp in length
* gap of unknown length
* 5164 6023: contig of 860 bp in length
* gap of unknown length
* 6024 6934: contig of 911 bp in length
* gap of unknown length
* 6935 7815: contig of 881 bp in length
* gap of unknown length
* 7816 8688: contig of 873 bp in length
* gap of unknown length
* 8689 9587: contig of 899 bp in length
* gap of unknown length
* 9588 10429: contig of 842 bp in length
* gap of unknown length
* 10430 11300: contig of 871 bp in length
* gap of unknown length
* 11301 12171: contig of 871 bp in length
* gap of unknown length
* 12172 13047: contig of 876 bp in length
* gap of unknown length
* 13048 13910: contig of 863 bp in length
* gap of unknown length
* 13911 14766: contig of 856 bp in length
* gap of unknown length
* 14767 15627: contig of 861 bp in length
* gap of unknown length
* 15628 16530: contig of 903 bp in length
* gap of unknown length
* 16531 17309: contig of 779 bp in length
* gap of unknown length
* 17310 18164: contig of 855 bp in length
* gap of unknown length
* 18165 19065: contig of 901 bp in length
* gap of unknown length
* 19066 19905: contig of 840 bp in length
* gap of unknown length
* 19906 20759: contig of 854 bp in length
* gap of unknown length
* 20760 21551: contig of 792 bp in length
* gap of unknown length
* 21552 22379: contig of 828 bp in length
* gap of unknown length
* 22380 23235: contig of 856 bp in length
* gap of unknown length
* 23236 24140: contig of 905 bp in length
* gap of unknown length
* 24141 25007: contig of 867 bp in length
* gap of unknown length
* 25008 25820: contig of 813 bp in length
* gap of unknown length
* 25821 26665: contig of 845 bp in length
* gap of unknown length
* 26666 27590: contig of 925 bp in length

* gap of unknown length
* 27591 28406: contig of 816 bp in length
* gap of unknown length
* 28407 29224: contig of 818 bp in length
* gap of unknown length
* 29225 30081: contig of 857 bp in length
* gap of unknown length
* 30082 30974: contig of 893 bp in length
* gap of unknown length
* 30975 31915: contig of 941 bp in length
* gap of unknown length
* 31916 32758: contig of 843 bp in length
* gap of unknown length
* 32759 33645: contig of 887 bp in length
* gap of unknown length
* 33646 34475: contig of 830 bp in length
* gap of unknown length
* 34476 35388: contig of 913 bp in length
* gap of unknown length
* 35389 36231: contig of 843 bp in length
* gap of unknown length
* 36232 37036: contig of 805 bp in length
* gap of unknown length
* 37037 37908: contig of 872 bp in length
* gap of unknown length
* 37909 38748: contig of 840 bp in length
* gap of unknown length
* 38749 39648: contig of 900 bp in length
* gap of unknown length
* 39649 40515: contig of 867 bp in length
* gap of unknown length
* 40516 41344: contig of 829 bp in length
* gap of unknown length
* 41345 42174: contig of 830 bp in length
* gap of unknown length
* 42175 43005: contig of 831 bp in length
* gap of unknown length
* 43006 43846: contig of 841 bp in length
* gap of unknown length
* 43847 44739: contig of 893 bp in length
* gap of unknown length
* 44740 45640: contig of 901 bp in length
* gap of unknown length
* 45641 46543: contig of 903 bp in length
* gap of unknown length
* 46544 47389: contig of 846 bp in length
* gap of unknown length
* 47390 48276: contig of 887 bp in length
* gap of unknown length
* 48277 49160: contig of 884 bp in length
* gap of unknown length
* 49161 50039: contig of 879 bp in length
* gap of unknown length
* 50040 50865: contig of 826 bp in length
* gap of unknown length
* 50866 51712: contig of 847 bp in length
* gap of unknown length
* 51713 52571: contig of 859 bp in length
* gap of unknown length
* 52572 53408: contig of 837 bp in length
* gap of unknown length
* 53409 54227: contig of 819 bp in length
* gap of unknown length
* 54228 55090: contig of 863 bp in length
* gap of unknown length
* 55091 55971: contig of 881 bp in length
* gap of unknown length
* 55972 56863: contig of 892 bp in length
* gap of unknown length
* 56864 57747: contig of 884 bp in length
* gap of unknown length
* 57748 58623: contig of 876 bp in length
* gap of unknown length

| | | | |
|---|-------|--------|-----------------------------|
| * | 58624 | 59480: | contig. of 857 bp in length |
| * | | gap of | unknown length |
| * | 59481 | 60365: | contig. of 885 bp in length |
| * | | gap of | unknown length |
| * | 60366 | 61266: | contig. of 901 bp in length |
| * | | gap of | unknown length |
| * | 61267 | 62134: | contig. of 868 bp in length |
| * | | gap of | unknown length |
| * | 62135 | 62991: | contig. of 857 bp in length |
| * | | gap of | unknown length |
| * | 62992 | 63853: | contig. of 862 bp in length |
| * | | gap of | unknown length |
| * | 63854 | 64832: | contig. of 979 bp in length |
| * | | gap of | unknown length |

```

* 23433 23532: gap of 100 bp
* 23533 24247: contig of 715 bp in length
* 24248 24347: gap of 100 bp
* 24348 25052: contig of 705 bp in length
* 25053 25152: gap of 100 bp
* 25153 25862: contig of 710 bp in length
* 25863 25962: gap of 100 bp
* 25963 26777: contig of 715 bp in length
* 26778 27505: contig of 728 bp in length
* 27506 27605: gap of 100 bp
* 27606 28341: contig of 736 bp in length
* 28342 28441: gap of 100 bp
* 28442 29147: contig of 706 bp in length
* 29148 29247: gap of 100 bp
* 29248 29896: contig of 649 bp in length
* 29897 29996: gap of 100 bp
* 29997 30720: contig of 724 bp in length
* 30721 30820: gap of 100 bp
* 30821 31536: contig of 716 bp in length
* 31537 31636: gap of 100 bp
* 31637 32362: contig of 726 bp in length
* 32363 32462: gap of 100 bp
* 32463 33158: contig of 666 bp in length
* 33159 33258: gap of 100 bp
* 33259 33964: contig of 706 bp in length
* 33965 34064: gap of 100 bp
* 34065 34788: contig of 724 bp in length
* 34789 35599: contig of 711 bp in length
* 35600 35699: gap of 100 bp
* 35700 36407: contig of 708 bp in length
* 36408 36507: gap of 100 bp
* 36508 37225: contig of 718 bp in length
* 37226 37325: gap of 100 bp
* 37326 38046: contig of 721 bp in length
* 38047 38146: gap of 100 bp
* 38147 38868: contig of 723 bp in length
* 38870 38969: gap of 100 bp
* 38970 39519: contig of 550 bp in length
* 39520 39619: gap of 100 bp
* 39620 40331: contig of 712 bp in length
* 40332 40431: gap of 100 bp
* 40432 41144: contig of 713 bp in length
* 41145 41244: gap of 100 bp
* 41245 41946: contig of 702 bp in length
* 41947 42046: gap of 100 bp
* 42047 42746: contig of 700 bp in length
* 42747 42846: gap of 100 bp
* 42847 43553: contig of 707 bp in length
* 43554 43653: gap of 100 bp
* 43654 44355: contig of 702 bp in length
* 44356 44455: gap of 100 bp
* 44456 45175: contig of 720 bp in length
* 45176 45275: gap of 100 bp
* 45276 45991: contig of 716 bp in length
* 45992 46091: gap of 100 bp
* 46092 46820: contig of 729 bp in length
* 46821 46920: gap of 100 bp
* 46921 47669: contig of 749 bp in length
* 47670 47769: gap of 100 bp
* 47770 48494: contig of 725 bp in length
* 48495 48594: gap of 100 bp
* 48595 49293: contig of 699 bp in length
* 49294 49393: gap of 100 bp
* 49394 50093: contig of 700 bp in length
* 50094 50193: gap of 100 bp
* 50194 50919: contig of 726 bp in length
* 50920 51019: gap of 100 bp
* 51020 51735: contig of 716 bp in length
* 51736 51835: gap of 100 bp
* 51836 52529: contig of 694 bp in length
* 52530 52629: gap of 100 bp

```

```

* 52630 53344: contig of 715 bp in length
* 53345 53444: gap of 100 bp
* 53445 54150: contig of 706 bp in length
* 54151 54250: gap of 100 bp
* 54251 54955: contig of 705 bp in length
* 54956 55055: gap of 100 bp
* 55056 55779: contig of 724 bp in length

Query Match      87.6% Score 18.4; DB 78; Length 64744;
Best Local Similarity 95.0% Pred. No. 91;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      2 gaagagaccatgaacaactt 21
Db      22249 GAAGATACCATGACAACTT 22268

RESULT      8
AL161719
LOCUS
DEFINITION
ACCESSION
AL161719
VERSION
AL161719.13 GI:13396426
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eulambda; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 215798)
REFERENCE
Submitted (13-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonequest@sanger.ac.uk
On Mar 20, 2001 this sequence version replaced gi:13379905.
COMMENT
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: ba3712
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: Plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 205502 bases at least Q40
Consensus quality: 211835 bases at least Q20
Insert size: 214196; sum-of-contigs
Quality coverage: 6.59x in Q20 bases; sum-of-contigs Quality
coverage: 11.54x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 20373 20472: contig of 20372 bp in length
* 20473 35049: contig of 14577 bp in length
* 35050 35149: gap of 100 bp
* 35150 126259: contig of 91110 bp in length
* 126260 126359: gap of 100 bp
* 126360 134145: contig of 7786 bp in length
* 134146 134245: gap of 100 bp
* 134246 175683: contig of 14148 bp in length
* 175684 175783: gap of 100 bp

```

```

* 175784 184379: contig of 8596 bp in length
* 184380 184479: gap of 100 bp
* 184480 188090: contig of 3611 bp in length
* 188091 188190: gap of 100 bp
* 188191 190453: contig of 2263 bp in length
* 190454 190553: gap of 100 bp
* 190554 193125: contig of 2572 bp in length
* 193126 193225: gap of 100 bp
* 193226 196254: contig of 3029 bp in length
* 196255 196354: gap of 100 bp
* 196355 198827: contig of 2473 bp in length
* 198828 198927: gap of 100 bp
* 198928 201913: contig of 2986 bp in length
* 201914 202013: gap of 100 bp
* 202014 204132: contig of 2119 bp in length
* 204133 204232: gap of 100 bp
* 204233 206775: contig of 2543 bp in length
* 206776 206875: gap of 100 bp
* 206876 209075: contig of 2200 bp in length
* 209076 209175: gap of 100 bp
* 209176 213317: contig of 4142 bp in length
* 213318 213417: gap of 100 bp
* 213418 215798: contig of 2381 bp in length.

```

```

FEATURES
Source
1. 215798
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone_id="RP11-37L2"
/clone_id="RP11-37L2"
1. 20372

```

```

misc_feature
/note="assembly_fragment:05293
fragment_chain:1"
20473..35049
/note="assembly_fragment:02239
fragment_chain:1"
35150..126259
/note="assembly_fragment:02452
fragment_chain:1"
126360..134145
/note="assembly_fragment:01565
fragment_chain:1"
134246..175683
/note="assembly_fragment:04916
fragment_chain:2"
175784..184379
/note="assembly_fragment:04426
fragment_chain:2"
184480..188090
/note="assembly_fragment:02922
fragment_chain:2"
188191..190453
/note="assembly_fragment:00003"
190554..193125
/note="assembly_fragment:00981"
193226..196254
/note="assembly_fragment:00991"
196355..198827
/note="assembly_fragment:01410"
198928..201913
/note="assembly_fragment:01589"
202014..204132
/note="assembly_fragment:02468"
204233..206775
/note="assembly_fragment:04019"
206876..209075
/note="assembly_fragment:04248"
209176..213317
/note="assembly_fragment:04549"
213418..215798
/note="assembly_fragment:04886"
215798..215798

```

```

BASE COUNT 65362 a 41441 c 41249 g 66128 t 1618 others
ORIGIN

```

```

Query Match 87.6%; Score 18.4; DB 79; Length 215798;
Best Local Similarity 95.0%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 2 gaagagaccatgaacactt 21
Db 27622 GAAGAGACCATGAACACTT 27641

```

```

RESULT 9
AC011647/c
LOCUS
DEFINITION
AC011647 147042 bp DNA HTG 20-SEP-2000
Homo sapiens clone RP11-15D18, WORKING DRAFT SEQUENCE, 7 unordered
pieces.
AC011647.5 GI:10198437
VERSION
HTG: HRGS_PHASE1; HRGS_DRAFT.
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 147042)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collamore, A.,
Cooke, P., Deatellano, K., Dewar, K., Domingo, M., Donelan, L., Doyle, M.,
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardy, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J., C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoczky, J., Lieu, C., Locke, K., MacDonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Melidri, J.,
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (08-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 20, 2000 this sequence version replaced gi:8096945.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RX/RepeatMasker.html

```

```

TITLE
JOURNAL
COMMENT

```

```

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 15.D.18
Center clone name: 15.D.18
----- Summary Statistics
Sequencing vector: M13; M77815; 97% of reads
Sequencing vector: Plasmid; n/a; 3% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 143564 bases at least Q40
Consensus quality: 145167 bases at least Q30
Consensus quality: 145878 bases at least Q20
Insert size: 143000; agarose-fp
Insert size: 146442; sum-of-contigs
Quality coverage: 7.6 in Q20 bases; agarose-fp
Quality coverage: 7.5 in Q20 bas.
NOTE: This is a 'working draft' sequence. It currently
consists of 7 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as

```

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 13561: contig of 13561 bp in length
* 13562 13661: gap of 100 bp
* 13662 14788: contig of 1127 bp in length
* 14789 14888: gap of 100 bp
* 14889 15930: contig of 1042 bp in length
* 15931 16030: gap of 100 bp
* 16031 39058: contig of 23029 bp in length
* 39060 39159: gap of 100 bp
* 39160 90776: contig of 51617 bp in length
* 90777 90876: gap of 100 bp
* 90877 119109: contig of 28233 bp in length
* 119110 119209: gap of 100 bp
* 119210 147042: contig of 27833 bp in length.
Location/Qualifiers
1. 147042
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-15D18"
/clone_1bp="RPC1-11 Human Male BAC"
1. 13561
/note="assembly-fragment
clone_end:SP6
vector_side:left"
misc_feature
13662..14788
/note="assembly-fragment"
misc_feature
14889..15930
/note="assembly-fragment"
misc_feature
16031..39059
/note="assembly-fragment"
misc_feature
39160..90776
/note="assembly-fragment"
misc_feature
90877..119109
/note="assembly-fragment"
misc_feature
119210..147042
/note="assembly-fragment
clone_end:T7
vector_side:right"

BASE COUNT 42217 a 29871 c 30225 g 44129 t 600 others
ORIGIN

Query Match 85.7%; Score 18; DB 62; Length 147042;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 aagagaccatgaacaact 20
|||||
Db 62098 AAGAGACCATGACACT 62081

RESULT 10
AC026083/c
LOCUS AC026083 19958 bp DNA HTG 01-SEP-2000
DEFINITION Homo sapiens chromosome 11 clone RP11-454022, WORKING DRAFT
SEQUENCE 19 unordered pieces.
ACCESSION AC026083
VERSION AC026083.3 GI:7658481
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 19958)
AUTHORS Waterston,R.H.
JOURNAL The sequence of Homo sapiens clone
REFERENCE 2 (bases 1 to 19958)
AUTHORS Waterston,R.H.
TITLE Direct Submission

JOURNAL
COMMENT

Submitted (19-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Apr 28, 2000 this sequence version replaced gi:7644580.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
Project Information
Center project name: H_NH0454022
----- Summary Statistics -----
Sequencing vector: M13; 1008
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 19185 bases at least Q40
Consensus quality: 194042 bases at least Q30
Consensus quality: 195420 bases at least Q20
Insert size: 187000; agarose-fp
Insert size: 198158; sum-of-contigs
Quality coverage: 4.82 in Q20 bases; agarose-fp
Quality coverage: 4.59 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1132: contig of 1132 bp in length
* 1133 1232: gap of unknown length
* 1233 2672: contig of 1440 bp in length
* 2673 2772: gap of unknown length
* 2773 3949: contig of 1177 bp in length
* 3950 4049: gap of unknown length
* 4050 6635: contig of 2586 bp in length
* 6636 6735: gap of unknown length
* 6736 9879: contig of 3144 bp in length
* 9880 9979: gap of unknown length
* 9980 12888: contig of 2909 bp in length
* 12889 12988: gap of unknown length
* 12989 16296: contig of 3308 bp in length
* 16297 16396: gap of unknown length
* 16397 20943: contig of 4547 bp in length
* 21044 21043: gap of unknown length
* 21044 27627: contig of 6584 bp in length
* 27628 27727: gap of unknown length
* 27728 35099: contig of 7372 bp in length
* 35100 35199: gap of unknown length
* 35200 44586: contig of 9387 bp in length
* 44587 44687: gap of unknown length
* 44688 52818: contig of 8132 bp in length
* 52819 52918: gap of unknown length
* 52919 63177: contig of 10259 bp in length
* 63178 63277: gap of unknown length
* 63278 74392: contig of 11115 bp in length
* 74393 74492: gap of unknown length
* 74493 90083: contig of 15591 bp in length
* 90084 90183: gap of unknown length
* 90184 108313: contig of 18130 bp in length
* 108314 108413: gap of unknown length
* 108414 132945: contig of 24532 bp in length
* 132946 133045: gap of unknown length
* 133046 161124: contig of 28079 bp in length
* 161125 161224: gap of unknown length
* 161225 199958: contig of 38734 bp in length.
Location/Qualifiers
1. 199958
/organism="Homo sapiens"

FEATURES
source


```
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-454022"
1. 1132
misc_feature
/note="assembly_name:Contig11"
1233. 2672
misc_feature
/note="assembly_name:Contig12"
2773. 3949
misc_feature
/note="assembly_name:Contig13"
4050. 6635
misc_feature
/note="assembly_name:Contig14
clone_end:r7
vector_side:right"
6736. 9879
misc_feature
/note="assembly_name:Contig15"
9980. 12888
misc_feature
/note="assembly_name:Contig16"
12989. 16296
misc_feature
/note="assembly_name:Contig17"
16397. 20943
misc_feature
/note="assembly_name:Contig18"
21044. 27627
misc_feature
/note="assembly_name:Contig19"
27728. 35099
misc_feature
/note="assembly_name:Contig20"
35200. 44586
misc_feature
/note="assembly_name:Contig21
clone_end:SP6
vector_side:left"
44687. 52818
misc_feature
/note="assembly_name:Contig22"
52919. 63177
misc_feature
/note="assembly_name:Contig23"
63278. 74392
misc_feature
/note="assembly_name:Contig24"
74493. 90083
misc_feature
/note="assembly_name:Contig25"
90184. 108313
misc_feature
/note="assembly_name:Contig26"
108414. 132945
misc_feature
/note="assembly_name:Contig27"
133046. 161124
misc_feature
/note="assembly_name:Contig28"
161225. 199958
misc_feature
/note="assembly_name:Contig29"
BASE COUNT 59966 a 39366 c 39871 g 58925 t 1810 others
ORIGIN
Query Match 85.7%; Score 18; DB 69; Length 19958;
Best Local Similarity 100.0%; Pred.No.1.6e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 3 aagagacatgacacact 20
|||||
Db 163010 AAGAGACCATGACACT 162993
```

```
RESULT 11
PST012352 2984 bp DNA BCT 10-MAY-1999
DEFINITION Pseudomonas stutzeri insertion sequence ISp1.
ACCESSION AJ012352
VERSION AJ012352.1 GI:4803706
KEYWORDS insertion sequence ISp1; ISL3 family; ORF1; ORF2; transposase.
SOURCE Pseudomonas stutzeri.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
Pseudomonas.
1 (bases 1 to 2984)
REFERENCE Bolognese,F., di Lecce,C., Galli,E. and Barbieri,P.
AUTHORS Activation and inactivation of pseudomonas stutzeri methylbenzene
TITLE catabolism pathways mediated by a transposable element
```

```
JOURNAL Appl. Environ. Microbiol. 65 (5), 1876-1882 (1999)
MEDLINE 99240403
REFERENCE 2 (bases 1 to 2984)
AUTHORS Barbieri,P.
TITLE Direct Submission
JOURNAL Submitted (29-OCT-1998) Barbieri P., Dept. Genetica e Biologia del
Microorganismi, University of Milan, Via Celoria 26, 20133 Milano,
ITALY
FEATURES
source
Location/Qualifiers
1. 2984
/organism="Pseudomonas stutzeri"
/insertion_seq="ISL3 family"
/strain="M1"
/db_xref="taxon:316"
/note="proposed name ISp1, inactivates catabolic genes"
complement(94. 492)
/note="ORF1"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB42635.1"
/db_xref="GI:4803707"
/translation="MDHEKVAASLAELGNSHRLSVREPLVAGHDGASVCDIQGLGI
PASTLSHLARMAKVLIRQEKHSRTVICIPEYGHLENLIGFLQECAGVRIAPHEPE
TALTQAQSLSSLSVSIAMKTAQKVTENMQ"
1649. 2926
CDS
/note="ORF2"
/codon_start=1
/transl_table=11
/product="putative transposase"
/protein_id="CAB42636.1"
/db_xref="GI:4803708"
/translation="MINSLSLPGVLYSDSDONEDIHRLTAPFPVACGGCGGEFY
RFGKNDVYRDLPHGKRVTLVAVRRRTTAGARHSCQPLDEWVGFRMLPLHEIV
EKESFNHPYTFVAQGTGDEKTVRDIENAKAEFLGRWRFETPRILGIDELYLNRVY
CILTNIERTLIDLATRRQDVYVNYLMLKLDROKVEIVSDMMNPYAAVAVLPQA
RIYVDFEYVIRMANDALERVRRGLRKLKPSQSRFLKQDRKILKRAHEVSDRELIM
ETWGAFPQLAAYEHRKRFYGIWDAATRIQAEALDWMATITIPGQKEWSDLVRAV
GNMKREMTYFETDMPYVNNATYESINRLAKNRGRGTSFVMAKRLYTTKKHKA
PFAKSPYKTKITGIGLPFAELNMGVDLSIT"
2961. 2984
/note="imperfect inverted repeats, 4 bp mismatch"
/rpt_unit=2961. 2984
/rpt_type=INVERTED
BASE COUNT 669 a 811 c 823 g 681 t
ORIGIN
Query Match 84.8%; Score 17.8; DB 3; Length 2984;
Best Local Similarity 90.5%; Pred.No.1.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 1 cgaagagacatgacacact 21
|||||
Db 2662 CGAAGAGACCATGACCTACTT 2682
```

```
RESULT 12
AF121970 6052 bp DNA BCT 07-MAY-1999
LOCUS AF121970/C
DEFINITION Pseudomonas aeruginosa topoisomerase (top), putative
transcriptional regulatory protein OhbR (ohbR), ortho-halobenzoate
1,2-dioxygenase beta-1S protein OhbA (ohbA), OhbC (ohbC),
ortho-halobenzoate 1,2-dioxygenase alpha-1S protein OhbB (ohbB),
and putative transposase A (tnpA) genes, complete cds.
ACCESSION AF121970
VERSION AF121970.1 GI:4406503
KEYWORDS
SOURCE Pseudomonas aeruginosa.
ORGANISM Pseudomonas aeruginosa.
Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
Pseudomonas.
1 (bases 1 to 6052)
```

AUTHORS Tsoi,T.V., Plotnikova,E.G., Cole,J.R., Guerlin,W.F., Bagdasarian,M.
and Tiedje,J.M.
TITLE Cloning, expression, and nucleotide sequence of the *Pseudomonas*
aeruginosa 142 ohb gene coding for oxygenolytic ortho
dehalogenation of halobenzoates
JOURNAL Appl. Environ. Microbiol. 65 (3), 2151-2162 (1999)
MEDLINE 99240444
REFERENCE 2 (bases 1 to 6052)
AUTHORS Hrymna,Y., Tsoi,T.V., Maltseva,O.V., Quensen III,J.F. and
Tiedje,J.M.
TITLE Construction and characterization of two recombinant bacteria that
grow on ortho- and para-substituted chlorobiphenyls
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 6052)
AUTHORS Tsoi,T.V. and Cole,J.R.
TITLE Direct Submission
JOURNAL Submitted (21-JAN-1999) Center for Microbial Ecology, Michigan
State University, A540 PSSB, East Lansing, MI 48824
FEATURES
source
1. 6052
/organism="Pseudomonas aeruginosa"
/strain="142"
/db_xref="taxon:287"
346..1563
/gene="top"
346..1563
/gene="top"
/note="top"
/codon_start=1
/transl_table=11
/product="topoisomerase"
/protein_id="A020003.1"
/db_xref="GI:4406504"
/translation="MRCRSYSLSCCKGKINHLRLPNNHTPTALSOQLDLDLSPGLADG
TKRAIAREYAIRGEGKROCAFYLIDGYIKDHPHVLINHLRIEHGGOVAIDHLI
NRMFWILSKSESLGFRISPGFEFYSONHYHAMPSTIEQARQRIEKLIDE
GLAPRLGYVMKPTIKTALIVDPKTYIRPDPOFTDSTYKADAFLESEKNEBAS
TYELGLAFRWAKGTIVDIGRKLAVHAFLEIDYRAKFGVQRAQRAKFAAIA
EREDYPERKQAGTGPSTPCPRCGKPMVLRTDSKRONGEOFGSGYPKCGIIPV
EVLVPAEDAGPVPTQOAIARAFAESPTDAPVCPKCGESVLMRSKRGSGKSRFWGC
PGFFPCRGMRSA6"
complement(1645..2361)
/gene="ohbR"
complement(1645..2361)
/gene="ohbR"
/note="IcIR family regulatory protein"
/codon_start=1
/transl_table=11
/product="putative transcriptional regulatory protein
OhbR"
/protein_id="A020004.1"
/db_xref="GI:4406505"
/translation="MSVKELADEVDMSTSRHRYLASLVKGLIERDAASGRYDLGEL
TQDGLAALNRLNPLVAEASKTLRLTVNDVDQTAFLVWVSERGPVIAHWHISSRPVIMN
LWNGSLPLITSSASRFLAFMPNPVLEATIDLSALALEPOAKKEFMELRQIAEV
QBDHLARVAGAVAGCLITASAPVLDAGFVAAMSLVATHTTMMERFEAIIPLCAGA
GVYSRLGTRPRPDPSAS"
2783..3313
/gene="ohbA"
2783..3313
/gene="ohbA"
/note="ortho-halobenzoate 1,2-dioxygenase terminal
oxygenase (ISP) small subunit"
/codon_start=1
/transl_table=11
/product="ortho-halobenzoate 1,2-dioxygenase beta-ISP
protein OhbA"
/protein_id="A020005.1"
/db_xref="GI:4406506"
/translation="NMTDSPKAIQPPADQOAVAVARQFEAEOLLVAVVEYIDDR
LEELPELFTEDGSGYQIISRENVGRGLPAPTMGCSNMLIDRIIVSRHANIVSOHAY
HVLNRYKAMEGDEALVQSNVYVMAIRGIGIARVYSTGKYVDRLVFEEDGPRFKQKR
VIFDQIDITLVLRPL"

gene complement(3249..4601)
/gene="ohbC"
CDS complement(3249..4601)
/gene="ohbC"
/codon_start=1
/transl_table=11
/product="ohbC"
/protein_id="A020007.1"
/db_xref="GI:4406508"
/translation="MEVGIOAORLVLPKALDRICRAIVHGARYLVDPHAHERTVL
VYDPRLNLHGAVYVQGVNAGDOINAGYARQFLVLYVDEALRPRLTLYPSNG
DDVSGREYTOQLHHQVREDKHHQIDVYVEFLALHDAVSGELRVYGVSLAEML
RGFLVGEVEFCQDQAPPREVETLRGLTGISPEBEGVQEMEDAGVAVGSRILGCOLP
VAVDALALKAEPQALAAVALDDLYHRHQVFLQRFHLIGAEAREDAVGVAGNQLGQA
VPALVEGLAHATLAAHAAQEGHAFELPLQVYVVLVNAAGVAAVALLTHORAAVRA
AVLHDDAAVLRSGDDGROPDMGGLVAVRLQRLQGHVGHPLSTENPLVIOFOIRL
VRNLVGHTEHRLFRRRGLGRYVGHMFSCSETLVRRNPALAGACACRLTLACRSRVFV
"
3354..4640
/gene="ohbB"
CDS 3354..4640
/gene="ohbB"
/note="ortho-halobenzoate 1,2-dioxygenase terminal
oxygenase (ISP) large subunit"
/codon_start=1
/transl_table=11
/product="ortho-halobenzoate 1,2-dioxygenase alpha-ISP
protein OhbB"
/protein_id="A020006.1"
/db_xref="GI:4406507"
/translation="MTNVSFQPTAEOALRPVYKFTTKAYLELENERIFRCKWAVY
ALEAEIPEPDYKATHIGLSPVITRAEDGSIHVMENRCARHRAALVCEORYGNASSLQ
CYVHOMDYNLKGLKCVPLRGVGRGKGMCKPDKSRHGLTKLRVATLNGIYVAFSSD
EYVPLEEYLCVPVAVELIKRCKRPLRLGFORIRISNMKLAENGRDYPHAGLLHF
LNFGEVLMNTSESASIOLOBRTGRLCLATELHTDEEAESAQAQAHVDEPTLRDSS
VNOSEOPDINISAVLSTFPRNLVYVQIDNALSTRVIRPGDEPFEMVWVRCGFVYDD
EELQRLRITHGILIGSGYVSLDEGAMERVDQSVSHDKDSTSPKMGVQDNTRSMYC
ATENPLKGEVQYIQTGLDADHFIEPQOASSGA"
complement(4731..5747)
/gene="tnpA"
/gene="tnpA"
complement(4731..5747)
/gene="tnpA"
/note="tnpA: similar to Serratia marcescens TnpA protein"
/codon_start=1
/transl_table=11
/product="putative transposase A"
/protein_id="A020008.1"
/db_xref="GI:4406509"
/translation="MWDGFRMTLRLEHYEKSEFNHPTFVNAOTGLDDEKTVDFINA
RAEFLGRHMRFEETPRILGIDELTLNKRKYCILNTERTLDDLATRRDDVYVNYLAK
LKDRQVVEIYSMDMNPYRAVAVLPLQRIATVDKRVHVRAMADALERKGLKLELQ
PSQSRVLKGDRIKLKRAHEVSDRERLIETWVGAPPOLLAADLAEHREYVGLWDAITR
LQAEALDELKATIPKQKEVMSDLRAVAGMNEEFMTVFEETMPATNAYTESINILA
KDNRRGRGVSFEVMARMLYTKHKKAPTAKVSPFYKKTIGYGLPDEELNLYCVD
LSTIR"
BASE COUNT 1220 a 1815 c 1849 g 1168 t
ORIGIN
Query Match 84.8%; Score 17.8; DB 1; Length 6052;
Best Local Similarity 90.5%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 cgaagagaccatgacacatt 21
|||||
Db 4995 CGAAGACACATGACCTACTT 4975
RESULT 13
AF027768 6921 bp DNA BCT 02-DEC-1998
LOCUS Serratia marcescens plasmid R471a RetA (retA), NucA (nucA), NucB
DEFINITION (mucB), TnpA (tnpA), and LspA (lspA) genes, complete cds.

```
ACCESSION   AF027768.1 GI:2688956
VERSION     AF027768.1
KEYWORDS
SOURCE      Serratia marcescens.
ORGANISM    Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Serratia marcescens.
REFERENCE   1 (bases 1 to 6921)
AUTHORS     Hedges,R.M., Rodriguez-Lemoine,V. and Datta,N.
TITLE       R factors from Serratia marcescens
JOURNAL     J. Gen. Microbiol. 86 (1), 88-92 (1975)
MEDLINE     75096689
REFERENCE   2 (bases 1 to 6921)
AUTHORS     Ho,C., Kulaeva,O.I., Levine,A.S. and Woodgate,R.
TITLE       A rapid method for cloning mutagenic DNA repair genes: isolation of
            unu-complementing genes from multidrug resistance plasmids R391,
            R46b, and R471a
JOURNAL     J. Bacteriol. 175 (17), 5411-5419 (1993)
MEDLINE     93374835
REFERENCE   3 (bases 1 to 6921)
AUTHORS     Kulaeva,O.I., Koonin,E.V., Wootton,J.C., Levine,A.S. and
            Woodgate,R.
TITLE       Unusual insertion element polymorphisms in the promoter and
            terminator regions of the mucAB-like genes of R471a and R446b
JOURNAL     Mutat. Res. 397 (2), 247-262 (1998)
MEDLINE     98202731
REFERENCE   4 (bases 1 to 6921)
AUTHORS     Woodgate,R.
TITLE       Direct Submission
JOURNAL     Submitted (02-OCT-1997) SDRM, NICHD/NIH, 9000 Rockville Pike,
            Bethesda, MD 20892-2725, USA
COMMENT     On Dec 16, 1997 this sequence version replaced gi:940364.
FEATURES
             location/qualifiers
             1..6921
             /organism="Serratia marcescens"
             /db_xref="taxon:615"
             /plasmid="R471a"
             /note="Incl/M R-plasmid"
             <1..538
             /note="hypothetical protein; similar to an hypothetical
             ort in R446b encoded by Genbank Accession Number AF027767"
             /codon_start=2
             /transl_table=1
             /protein_id="AAC82520.1"
             /db_xref="GI:2688958"
             /translation="KNTLFTHGAVLYLVCFLVYRWRVAPLPVGCGLRVGAPWLY
             GATAVYVGEFLGKVLDDPAEPKTLFADKSLPDVTLTLAFLIAPLNEETLFGS
             IMLNFSRYSWTMMGLALITSLFVAHSOYOMLTLIAELFLVGLITSARIRSGGL
             LPLVILHMEATAGLLEFG"
             complement(715..2202)
             /gene="reta"
             complement(715..2202)
             /gene="reta"
             /codon_start=1
             /transl_table=1
             /product="Reta"
             /db_xref="GI:2688957"
             /protein_id="AAC82519.1"
             /translation="MMTGEKSDSLVIAAKOANPKGAESVERRSGAKNAEQPMHRR
             TQRESMSRLSRVREAKOKKKEFTALPHLLVEALEAFLSLSRKAAGVGIWM
             MDYAGNMKNNTIDLRHLHOGSYRAQGRHRYIPKADGNORPLIASLEIKIVQALY
             KILNAVYENDFMGSGYFGRPSOHDALDALATGLVTRNVMVLADADISOPEDRVSHE
             WLIRFPHRTGDRRVIRLIRKMLTACSECGMAREGTGQAVISPLANILHYV
             FDIAMHQRRRRYATGNVYRYADDIYIGDKRDARFRIRAKORLREGLVHPEK
             TLRMEGRFAENRAIRKGKPEFTNGLFTTHSGKORNGRFMLIRTRTRDMRTATLK
             AIKDGLRRNHYSIPEQGWLRVVOGYLNVHSVGNFPTMOKFRTHVTNIMRRALRR
             RSQKDDTWTKANKLAAMLPRVFLHPWEVERPTAHROEPGA"
             2539..2544
             2563..2568
             2566..2583
             /bound_moiety="LexA"
             /product="LexA"
             2588..2592
             /gene="mucA"
             2588..3034
             /gene="mucA"
             2600..3034
             /gene="mucA"
             /note="similar to MucA from plasmid R446b, Genbank
             Accession Number AF027767; also similar to Escherichia
             coli Umud, Swiss-Prot Accession Number P04153, and to
             Salmonella typhimurium Umud, Swiss-Prot Accession Number
             P22493, Samh, Swiss-Prot Accession Number P23831, MucA,
             Swiss-Prot Accession Number P07376, and ImpA, Swiss-Prot
             Accession Number P18641"
             /codon_start=1
             /transl_table=1
             /product="MucA"
             /protein_id="AAC82521.1"
             /db_xref="GI:2688959"
             /translation="MALDLVYPRPSVEGDALPFFAEPYAGFPSPAGYEANELNLD
             YCVRHPAATYFLRVSGDSMDRARIHGDVVLVDRDAEERHSGIYASIDNEFTVKQLQ
             LRPPCLMPILNPAFPPIREFDEALQIMGVVFVYMRHRCPTP"
             3022..4287
             /gene="mucB"
             3022..4287
             /gene="mucB"
             /note="similar to Muc from plasmid R446b, Genbank
             Accession Number AF027767; also similar to Escherichia
             coli Umuc, Swiss-Prot Accession Number P04152, and to
             Salmonella typhimurium Umuc, Swiss-Prot Accession Number
             P22494, Samh, Swiss-Prot Accession Number P23832, MucB,
             Swiss-Prot Accession Number P14303, and ImpB, Swiss-Prot
             Accession Number P18642"
             /codon_start=1
             /transl_table=1
             /product="MucB"
             /protein_id="AAC82522.1"
             /db_xref="GI:2688960"
             /translation="MYALIDVNSMYSCSEAFRPDLRGRPVVLSNNDGALVAVNRRA
             KALGVRRGEPEYRCRLRLEOHNAVFNSSNYTLVASFSAFATVSNESLAPRFVYSIDE
             LFADASHMEGVSPQARFQGLLRAEVLRTITLTCGVGVAPTRTLAKICNHAKTAPAG
             GVALASDPYRLKRLMSLYSAGDVWGVGRTREKGLATMGIKTKALALAEADITLARLIG
             VTLERTIRLELNGKAFLEBGTGAKQDLVRSRSGSRPTOLAQMOQAVTKYATAGKE
             LRQENRLAKVTVFIRSSAYDAGVPSYNOATEVLSVPSQPTROIITAAOGLSRIMR
             EGVRYAKAGVMLSDQGHETQDLFAPAAVPPGSEKIMATIDKINREGSRVFPAGEG
             IDPEYAMRREMLSPAYTTQMKDIPAAFR"
             4315..6097
             /organism="Serratia marcescens"
             /db_xref="taxon:615"
             /insertion_seq="IS1396"
             complement(4830..6044)
             /gene="tnpa"
             complement(4830..6044)
             /gene="tnpa"
             /note="putative transposase"
             /codon_start=1
             /transl_table=1
             /product="tnpa"
             /protein_id="AAC82523.1"
             /db_xref="GI:2688961"
             /translation="MHFQVDPDPIACEECGVGEFVRFGKRDVYRDLPIHGRVTL
             MWRRRYTCRACKTTFRQVLPDEWVDGFMRLTLREYVEKESFNHPYFVAAGTGLDEK
             TVRDIFNARAFELTRMHRFEPTILIGIDELINKRYCIIITNIBERTLIDLATRRDQ
             VTDVYILAKIKROKVEIYFMSMDMMPYRAVAVAPORIVYDKRHVYBMADALERYR
             KGLRKEKLPQSRLTGDRIKILKRAHVSRELIINMETWNGARPOLLANENHERRY
             GIMWATTRLQAEALDENIATIPRGQKEWSDLVRAVGNREETWTFETDMPVTNAY
             TESTINRLAKDNRNREGRGVSFPMVARMRLYTTKHKKAPJAKVAVSPFYKKTGYGLDPA
             BELVYGVDSLTI"
             complement(6141..6653)
             /gene="lspsa"
             complement(6141..6653)
             /gene="lspsa"
             /codon_start=1
             /transl_table=1
             /product="lspsa"
             /protein_id="AAC82524.1"
```


TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (11-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 22, 2000 this sequence version replaced gi:8195533.
All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: W1BR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project information

Center project name: L8872

Center clone name: 186.A.21

Summary Statistics

Sequencing vector: M13; M77815; 54% of reads

Sequencing vector: Plasmid; n/a; 46% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 200951 bases at least Q40

Consensus quality: 204253 bases at least Q30

Insert size: 198000; agarose-fp

Insert size: 204638; sum-of-contigs

Quality coverage: 6.7 in Q20 bases; agarose-fp

Quality coverage: 6.5 in Q20 b.

NOTE: This is a 'working draft' sequence. It currently

consists of 11 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 4471: contig of 4471 bp in length

4472 4571: gap of 100 bp

4572 4645: contig of 74 bp in length

4646 4745: gap of 100 bp

4746 10308: contig of 5563 bp in length

10309 10408: gap of 100 bp

10409 64985: contig of 54577 bp in length

64986 65085: gap of 100 bp

65086 78435: contig of 13350 bp in length

78436 78535: gap of 100 bp

78536 93848: contig of 15313 bp in length

93849 93948: gap of 100 bp

93949 109475: contig of 15527 bp in length

109476 109575: gap of 100 bp

109576 130155: contig of 20580 bp in length

130156 130255: gap of 100 bp

130256 158585: contig of 28330 bp in length

158586 158685: gap of 100 bp

158686 201676: contig of 42991 bp in length

201677 201776: gap of 100 bp

201777 205638: contig of 3862 bp in length.

Location/Qualifiers

1. 205638

/organism="Mus musculus"

/db_xref="taxon:10090"

/chromosome="5"

/map="5"

/clone="RP23-186A21"

/clone.lib="RPC1-23 Female Mouse BAC"

1. 4471

/note="assembly_fragment"

clone_end:SP6

vector_side:left

4572. 4645

/note="assembly_fragment"

4746. 10308

/note="assembly_fragment"

10409. 64985

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature /note="assembly_fragment"
65086. 78435
misc_feature /note="assembly_fragment"
78536. 93848
misc_feature /note="assembly_fragment"
93949. 109475
misc_feature /note="assembly_fragment"
109576. 130155
misc_feature /note="assembly_fragment"
130256. 158585
misc_feature /note="assembly_fragment"
158686. 201676
misc_feature /note="assembly_fragment"
201777. 205638
misc_feature /note="assembly_fragment"
clone_end:T7
vector_side:right"
BASE COUNT 64156 a 38300 c 37657 g 64521 t 1004 others
ORIGIN

Query Match 84.8%; Score 17.8; DB 71; Length 205638;
Best Local Similarity 90.5%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CGAAGACCATGACCACTT 21
DB 133856 CGAAGACCATGAAAAACAT 133836

Search completed: September 21, 2001, 00:36:50
Job time: 29665 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 02:42:46 : Search time 330.77 Seconds
(without alignments)
39.864 Million cell updates/sec

Title: US-09-138-735-12

Perfect score: 21

Sequence: 1 cgaagagaccatgaacaactt 21

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_0601.*
1: /cgnl_9/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /cgnl_9/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /cgnl_9/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /cgnl_9/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /cgnl_9/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /cgnl_9/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /cgnl_9/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /cgnl_9/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /cgnl_9/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /cgnl_9/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /cgnl_9/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /cgnl_9/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /cgnl_9/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /cgnl_9/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /cgnl_9/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /cgnl_9/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /cgnl_9/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /cgnl_9/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /cgnl_9/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /cgnl_9/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /cgnl_9/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /cgnl_9/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|-----------------------|
| 1 | 21 | 100.0 | 21 | AAT27315 | T. cruzi epimastigote |
| 2 | 21 | 100.0 | 21 | AAK84097 | PCR primer for T. |
| 3 | 21 | 100.0 | 3402 | AAT27310 | Trypanosoma cruzi |
| 4 | 21 | 100.0 | 3402 | AAK84092 | T. cruzi P1040 cod |
| 5 | 16.8 | 80.0 | 747 | AAAO1873 | Human colon cancer |
| 6 | 16.8 | 80.0 | 1606 | AAAC64747 | Secreted protein K |
| 7 | 16.8 | 80.0 | 1931 | AAAC98922 | Human pancreatic C |
| 8 | 16.8 | 80.0 | 2163 | AAAF44684 | Novel protein kina |
| 9 | 16.8 | 80.0 | 2241 | AAVS9121 | Nucleotide sequenc |
| 10 | 15.8 | 75.2 | 1815 | AAO42168 | ODC cDNA. Synthet |
| 11 | 15.8 | 75.2 | 1958 | AAFI8267 | Lung cancer associ |

| | | | | | | |
|----|------|------|------|----|-----------|-----------------------|
| 12 | 15.8 | 75.2 | 2265 | 21 | AA244288 | Human brain PABP3 |
| 13 | 15.8 | 75.2 | 2265 | 21 | AA244289 | Human uterus type |
| 14 | 15.4 | 73.3 | 1913 | 21 | AAAC64441 | Human secreted pro |
| 15 | 15.4 | 73.3 | 1915 | 11 | AAQ04775 | Fragment contig. Ch |
| 16 | 15.4 | 73.3 | 1980 | 18 | AAT59454 | H-Delta-1 contig c |
| 17 | 15.4 | 73.3 | 2135 | 21 | AAAC65395 | Human carbohydrate |
| 18 | 15.4 | 73.3 | 2533 | 19 | AAT96031 | Mouse gamma II ada |
| 19 | 15.4 | 73.3 | 2663 | 18 | AAT70174 | Proliferation and |
| 20 | 15.4 | 73.3 | 2663 | 20 | AAI16300 | Human delta-1 prot |
| 21 | 15.4 | 73.3 | 2683 | 18 | AAI16817 | Human delta-1 gene |
| 22 | 15.4 | 73.3 | 2883 | 18 | AAT58898 | C-Delta-1 gene (al |
| 23 | 15.4 | 73.3 | 2933 | 21 | AAAS4105 | PRO172 cDNA. Homo |
| 24 | 15.4 | 73.3 | 2933 | 21 | AAAC58587 | Human PRO172 prote |
| 25 | 15.4 | 73.3 | 2933 | 21 | AAAT7512 | Human PRO172 cDNA |
| 26 | 15.4 | 73.3 | 2933 | 21 | AAZ93705 | PRO172 DNA35916-11 |
| 27 | 15.4 | 73.3 | 2933 | 22 | AAAC97368 | Human angiotensin |
| 28 | 15.4 | 73.3 | 3162 | 21 | AAZ98679 | Human delta protei |
| 29 | 15.2 | 72.4 | 359 | 21 | AAI12677 | PCR primer PEXMYC2 |
| 30 | 15.2 | 72.4 | 408 | 22 | AAZ79917 | Human colon cancer |
| 31 | 15.2 | 72.4 | 504 | 20 | AAZ25115 | Novel human polynu |
| 32 | 15.2 | 72.4 | 504 | 20 | AAZ25115 | wheat anthranilate |
| 33 | 15.2 | 72.4 | 675 | 21 | AAAC3972 | Arabidopsis thalia |
| 34 | 15.2 | 72.4 | 915 | 13 | AAO31798 | PRP4 gene. Nicotia |
| 35 | 15.2 | 72.4 | 1029 | 21 | AAAC75875 | Human ORFX ORF130 |
| 36 | 15.2 | 72.4 | 1046 | 12 | AAO10377 | Plasmodium PMG3C9 use |
| 37 | 15.2 | 72.4 | 1056 | 21 | AAAC55832 | S. lavendulae Mmcr |
| 38 | 15.2 | 72.4 | 1221 | 21 | AAAC9196 | Arabidopsis thalia |
| 39 | 15.2 | 72.4 | 1221 | 21 | AAAC40196 | Arabidopsis thalia |
| 40 | 15.2 | 72.4 | 1274 | 21 | AAAC45280 | Arabidopsis thalia |
| 41 | 15.2 | 72.4 | 1371 | 21 | AAAC44690 | Arabidopsis thalia |
| 42 | 15.2 | 72.4 | 1393 | 21 | AAAC46421 | Arabidopsis thalia |
| 43 | 15.2 | 72.4 | 1395 | 21 | AAAC36574 | Arabidopsis thalia |
| 44 | 15.2 | 72.4 | 1614 | 21 | AAAC45832 | Arabidopsis thalia |
| 45 | 15.2 | 72.4 | 1624 | 21 | AAZ65339 | Human secreted pro |

ALIGNMENTS

| | |
|--|------------------------|
| RESULT 1 | |
| ID AAT27315 | standard; cDNA: 21 BP. |
| XX AAT27315; | |
| XX 26-NOV-1996 | (first entry) |
| DT T. cruzi epimastigote P1040t antigen primer corresp. | to bases 1997-2017. |
| DE Antigen: Trypanosoma cruzi; epimastigote; serum; Chagas disease; probe; | |
| XX Primer: PCR; polymerase chain reaction; amplification; antibody; ss. | |
| KW Synthetic. | |
| XX FR2723589-A1. | |
| XX 16-FEB-1996. | |
| PD 12-AUG-1994; | 94FR-0010132. |
| XX 12-AUG-1994; | 94FR-0010132. |
| PR (IMMR) BIO MERIEUX. | |
| PA Jollivet M, Lesenechal M, Paranhos-Baccala G; | |
| XX WPI; 1996-190287/20. | |
| DR New nucleic acid encoding Trypanosoma cruzi epimastigote antigen - | |
| XX FT useful for diagnosis, monitoring and therapy of Chagas disease | |
| XX Claim 29; Page 40; 55pp; French. | |

CC The primers AAT27311-5 were used to PCR amplify the sequence encoding a
 CC novel isolated antigenic protein from Trypanosoma cruzi epimastigotes,
 CC designated PTC100t (AAT27310). The primers, derived from the sequences
 CC of a 594 and 1041 bp fragment of PTC100t, amplified the gene as 3
 CC fragments. This primer corresponds to nucleotides 1997-2017 of the PTC100t
 CC sequence, derived from the 1041 bp fragment. The 594 bp fragment was
 CC isolated from a T. cruzi genomic expression library in lambda gt11, using
 CC a mixture of sera from patients with Chagas disease. It corresponds to
 CC nucleotides 1232-1825 of PTC100t. The 1041 bp fragment was isolated from
 CC a lambda gt10 library using the 594 bp fragment as a probe. The protein,
 CC or antibodies raised against it, can be used in the detection and
 CC monitoring of T. cruzi infection i.e. Chagas disease.
 XX
 SO Sequence 21 BP; 9 A; 5 C; 4 G; 3 T; 0 other;

Query Match 100.0%; Score 21; DB 17; Length 21;
 Best Local Similarity 100.0%; Pred. No. 0.3;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cgaagagaccatgaacaactt 21
 ||||||||||||||||||
 Db 1 cgaagagaccatgaacaactt 21

RESULT 2

AAX84097
 ID AAX84097 standard; DNA; 21 BP.

AC AAX84097;

XX 27-AUG-1999 (first entry)

DE PCR primer for T. cruzi PTC40 coding sequence.

XX PTC40; Tc40; infection; diagnosis; immune complex; antigenic determinant;
 KM therapy; antibody; PCR primer; ss.

OS Synthetic.

OS Trypanosoma cruzi.

XX WO929867-A1.

XX 17-JUN-1999.

XX 10-DEC-1998; 98WO-IB01987.

XX 10-DEC-1997; 97US-0988242.

XX (INMR) BIO MERIEUX.

XX Jolivet M, Lesenechal M, Mandrand B, Paranhos-Baccala G;

XX WPI; 1999-394978/33.

XX New Trypanosoma cruzi antigen

PS Disclosure; Page 22; 65pp; English.

XX This sequence is a PCR primer for DNA encoding the Trypanosoma cruzi
 CC PTC40 protein of the invention. The PTC40 antigenic determinant is
 CC useful as a reagent for detection and/or monitoring of Trypanosoma cruzi
 CC infection from samples including blood serum or plasma, urine, saliva, or
 CC tears, by contacting with the sample and detecting an immune complex. The
 CC PTC40 antigenic determinant, the vector, expression cassette, cell or
 CC antibody are useful for treatment or prevention (vaccine) of a
 CC Trypanosoma cruzi infection in a man or animal. Current Trypanosoma cruzi
 CC antigens are obtained from protein fractions of the noninfectious stage
 CC of the parasite, and these do not allow sufficient production of antigens
 CC for use in reliable serological diagnostic tests. The strain to strain
 CC polymorphism reduces reliability of the tests.
 XX
 SO Sequence 21 BP; 9 A; 5 C; 4 G; 3 T; 0 other;

Query Match 100.0%; Score 21; DB 20; Length 21;
 Best Local Similarity 100.0%; Pred. No. 0.3;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cgaagagaccatgaacaactt 21
 ||||||||||||||||||
 Db 1 cgaagagaccatgaacaactt 21

RESULT 3

AAT27310
 ID AAT27310 standard; cDNA; 3402 BP.

AC AAT27310;

XX 26-NOV-1996 (first entry)

DE Trypanosoma cruzi epimastigotic PTC100t antigen gene.

XX Antigen; Trypanosoma cruzi; epimastigote; serum; Chagas disease; probe;

KM Primer; PCR; polymerase chain reaction; amplification; antibody; ds.

XX Trypanosoma cruzi.

XX Key Location/Qualifiers

FT CDS 266..3013 /tag= a

FT /product= PTC100t epimastigotic antigen

PN FR273589-A1.

XX 16-FEB-1996.

XX 12-AUG-1994; 94FR-0010132.

XX 12-AUG-1994; 94FR-0010132.

XX (INMR) BIO MERIEUX.

XX Jolivet M, Lesenechal M, Paranhos-Baccala G;

XX WPI; 1996-190287/20.

XX P-PsDB; AAR91615.

XX New nucleic acid encoding Trypanosoma cruzi epimastigotic antigen -
 XX useful for diagnosis, monitoring and therapy of Chagas disease

XX Claim 1; Page 24-26; 55pp; French.

XX This is the nucleotide sequence encoding a novel isolated antigenic
 CC protein from Trypanosoma cruzi epimastigotes, designated PTC100t.
 CC The clone Tc50 was isolated from a T. cruzi genomic expression library in
 CC lambda gt11, using a mixture of sera from patients with Chagas disease.
 CC Clone Tc50 contained an 594 bp insert corresp. to nucleotides 1232-1825
 CC of this sequence. The Tc50 sequence was subsequently used to probe a
 CC Southern blot of restriction enzyme digested T. cruzi DNA and also screen
 CC a lambda gt10 library to isolate a 1041 bp EcoRI fragment corresp. to
 CC nucleotides 1403-2443 of PTC100t. Primers (AAT27311-5) were synthesised
 CC based on the sequences of the 594 and 1041 bp fragments and used to
 CC amplify the PTC100t clone as 3 fragments from cDNA derived from mRNA
 CC purified from T. cruzi epimastigotes. The protein or antibodies raised
 CC against it can be used in the detection and monitoring of T. cruzi
 CC infection i.e. Chagas disease.
 XX

SO Sequence 3402 BP; 889 A; 818 C; 958 G; 737 T; 0 other;

Query Match 100.0%; Score 21; DB 17; Length 3402;
 Best Local Similarity 100.0%; Pred. No. 0.61;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cgaagagacatgaacaactt 21
|||||
Db 1997 cgaagagacatgaacaactt 2017

RESULT 4

AAx84092
ID AAX84092 standard; cDNA; 3402 BP.

XX AAX84092;

XX 27-AUG-1999 (first entry)

DE T. cruzi PTC40 coding sequence.

XX PTC40; Tc40; infection; diagnosis; immune complex; antigenic determinant;
KW therapy; antibody; ds.

XX Trypanosoma cruzi.

XX WO9929867-A1.

XX 17-JUN-1999.

XX 10-DEC-1998; 98WO-IB01987.

XX 10-DEC-1997; 97US-0988242.

XX (IMMR) BIO MERIEUX.

PI Jolivet M, Lesenechal M, Mandrand B, Paranhos-Baccala G;

DR WPI: 1999-394978/73.

DR P-PSDB: AAY22124.

XX New Trypanosoma cruzi antigen

PS Claim 1; Page 52-56; 65pp; English.

XX This sequence encodes the Trypanosoma cruzi PTC40 protein of the
CC invention, and is designated Tc40. The PTC40 antigenic determinant is
CC useful as a reagent for detection and/or monitoring of Trypanosoma cruzi
CC infection from samples including blood serum or plasma, urine, saliva, or
CC tears, by contacting with the sample and detecting an immune complex. The
CC PTC40 antigenic determinant, the vector, expression cassette, cell or
CC antibody are useful for treatment or prevention (vaccine) of a
CC Trypanosoma cruzi infection in a man or animal. Current Trypanosoma cruzi
CC antigens are obtained from protein fractions of the noninfectious stage
CC of the parasite, and these do not allow sufficient production of antigens
CC for use in reliable serological diagnostic tests. The strain to strain
CC polymorphism reduces reliability of the tests.

XX Sequence 3402 BP; 888 A; 821 C; 956 G; 737 T; 0 other;

Query Match 100.0%; Score 21; DB 20; Length 3402;

Best Local Similarity 100.0%; Pred. No. 0.61; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cgaagagacatgaacaactt 21
|||||
Db 1997 cgaagagacatgaacaactt 2017

RESULT 5

AAA01873
ID AAA01873 standard; cDNA; 747 BP.

XX AAA01873;

XX 19-MAY-2000 (first entry)

XX Human colon cancer cell line polynucleotide sequence SEQ ID NO:1864.

XX Human: colon cancer; tumour; diagnosis; gene expression product;
KW probe; detection; cancerous state; metastasis; identification;
KW breast cancer; oestrogen receptor-positive breast cancer; therapy;
KW oestrogen receptor-negative breast cancer; lung cancer; ss.

XX Homo sapiens.

XX WO9958675-A2.

XX 18-NOV-1999.

XX 13-MAY-1999; 99WO-US10602.

XX 14-MAY-1998; 98US-0085426.

XX 15-MAY-1998; 98US-0085537.

XX 21-OCT-1998; 98US-0105234.

XX 27-OCT-1998; 98US-0105877.

XX (CHIR) CHIRON CORP.

XX (HSE-) HSE INC.

PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;

PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;

PI Lanson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;

PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;

DR WPI: 2000-126369/11.

XX Polynucleotide library used to determine cancerous states of mammalian

XX cells -

PS Claim 1; Page 697-698; 1097pp; English.

XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
CC libraries constructed from human colon cancer cell lines. The present
CC invention also describes a method of detecting differentially expressed
CC genes correlated with a cancerous state of a mammalian cell, comprising
CC detecting at least one differentially expressed gene product in a test
CC sample derived from a cell suspected of being cancerous, where detection
CC of the differentially expressed gene product is correlated with a
CC cancerous state of the cell from which the test sample was derived.
CC The polynucleotide sequences can be used in a method for detecting
CC differentially expressed genes correlated with a cancerous state of a
CC mammalian cell. The polynucleotides can also be used as probes for
CC detecting and mapping related genes. They can be used in diagnosis and
CC prognosis of diseases and disorders (e.g. identification of
CC pre-metastatic or metastatic cancerous states, stages of cancer, or
CC responsiveness of cancer to therapy). This is particularly for breast
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC negative breast cancer, lung cancer, and colon cancer.

XX Sequence 747 BP; 136 A; 227 C; 193 G; 179 T; 12 other;

Query Match 80.0%; Score 16.8; DB 21; Length 747;

Best Local Similarity 90.0%; Pred. No. 50; Mismatches 2; Indels 0; Gaps 0;

Qy 2 gaagagacatgaacaactt 21
|||||
Db 341 gaagagacatgaacaactt 360

RESULT 6

AAC64747
ID AAC64747 standard; cDNA; 1606 BP.

XX AAC64747;

XX 28-FEB-2001 (first entry)

XX

| | | |
|--------------------------|--|--|
| DE | | Secreted protein K139_11 encoding cDNA SEQ ID NO:5. |
| xx | | |
| KW | | Secreted protein; AM179_11; BG221_11; K139_11; K511_11; N154_11; |
| KM | | immunomodulator; antiinflammatory; immunosuppressive; cyostatic; * |
| KM | | vulnerary; osteopathic; antiaarthritic; neuroprotective; antineumatic; |
| KM | | antiparkinsonian; nootropic; analgesic; vasotrophic; cerebroprotective; |
| KM | | contraceptive; haemostatic; thrombolytic; antibacterial; nutritional; |
| KM | | cytokine; cell proliferation; cell differentiation; immunomodulatory; |
| KV | | immune disorder; proliferative disorder; regulation; haematopoiesis; ss. |
| xx | | |
| OS | Mus musculus. | |
| XN | | |
| PN | WO2006148-A1. | |
| PD | 09-NOV-2000. | |
| XX | | |
| PX | 29-APR-2000; 2000WO-US11818. | |
| PE | | |
| PR | 29-APR-1999; 99US-0131596. | |
| PA | (GENY) GENETICS INST INC. | |
| XX | | |
| PI | Jacobs K, McCoy JM, Lavallic ER, Collins-Racie LA, Evans C; | |
| PT | Merberg D; | |
| DR | WPt: 2001-015913/02. | |
| XX | P-PStDB; AAB36408. | |
| PT | Noval AW17911, BG22111, K13911 and N15411 secreted proteins, | |
| PS | their fragments and polynucleotides, useful in diagnostic and research | |
| PT | assays and for treating, e.g. immune and proliferative disorders and | |
| PT | for regulating hematopoiesis - | |
| PS | Claim 9; Page 60; 67Pp; English. | |
| CC | The present invention describes secreted proteins designated AW179_11, | |
| CC | BG221_11, K139_11, K511_11 and N154_11. The secreted proteins can have | |
| CC | immunomodulator, antiinflammatory, immunosuppressive, cyostatic, | |
| CC | vulnerary, osteopathic, antiarthritis, neuroprotective, antirheumatic, | |
| CC | antiparkinsonian, nootropic, analgesic, vasotrophic, cerebroprotective, | |
| CC | contraceptive, haemostatic, thrombolytic and antibacterial activities. | |
| CC | The proteins can be used in assays to determine biological activity and | |
| CC | as markers for tissues in which the corresponding protein is | |
| CC | preferentially expressed (either constitutively or at a particular stage | |
| CC | of tissue differentiation or development or in a disease state). The | |
| CC | proteins can be used as nutritional sources or supplements and may | |
| CC | exhibit cytokine, cell proliferation, cell differentiation and | |
| CC | immunomodulatory activities. They may be useful in the treatment of | |
| CC | immune and proliferative disorders and for regulating haematopoiesis. | |
| CC | The present sequence encodes the K139_11 secreted protein from the | |
| CC | present invention. | |
| SO | | |
| XX | Sequence 1606 BP; 389 A; 448 C; 412 G; 357 T; 0 other: | |
| Query Match | 80.0%; Score 16.8; DB 22; Length 1606; | |
| Best Local Similarity | 90.0%; Pred. NO. 56; | |
| Matches 18; Conservative | 0; Mismatches 2; Indels 0; Gaps 0; | |
| Oy | 2 gaagagaccatgaacaactc 21 | |
| Db | 1068 gaagaaccttgacaacglt 1087 | |
| RESULT 7 | | |
| AAC98922 | ID AAC98922 standard; CDNA; 1931 BP. | |
| XX | | |
| AC | AAC98922; | |
| DT | 09-MAR-2001 (first entry) | |
| XX | | |
| DE | Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:150. | |

XX Human: pancreas; pancreatic cancer; pancreatic cancer antigen;
 KW detection; diagnosis; identification; cytosolic; neuroprotective;
 KW nontoxic; immunomodulatory; relaxant; contraceptive; gynaecological;
 KW antinflammatory; cardiac; gene therapy; chromosome mapping;
 KW linkage analysis; tissue identification; tissue typing; forensic;
 KW neural; immune system; muscular; reproductive; gastrointestinal;
 KW pulmonary; cardiovascular; renal; proliferative; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO20055320-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05989.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 P1 WPI: 2000-579444/54.
 XX
 DR P-PSDB: AAB54457.
 XX
 PT New nucleic acid that is a pancreatic cancer antigen for preventing,
 PT treating, or ameliorating a medical condition, particular pancreatic
 PT cancer, or for use in assays for diagnosing a pathological condition -
 PT
 PS Claim 1; Page 608-609; 1379pp; English.
 XX
 XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
 CC proteins, called pancreatic cancer antigens, given in AAB54008 to
 CC AAB54466. The human pancreatic cancer antigens have cytosolic,
 CC neuroprotective, nontoxic, immunomodulatory, relaxant, contraceptive,
 CC gynaecological, cardiac and antinflammatory activities, and can be used
 CC in gene therapy. The polynucleotide and proteins can be used for
 CC preventing, treating, or ameliorating a medical condition or in assays
 CC for diagnosing a pathological condition or a susceptibility to one in a
 CC subject. Binding partners to the proteins and the activity of the
 CC proteins can be identified. The pancreatic cancer antigens can be used to
 CC detect, treat or prevent pancreatic disorders, especially cancer.
 CC Agonists and antagonists to the antigens can be screened for. The
 CC pancreatic cancer antigen polynucleotides can be used to design nucleic
 CC acid hybridisation probes that can be used in chromosome mapping, linkage
 CC analysis, tissue identification and/or typing and a variety of forensic
 CC and diagnostic methods. The proteins can be used to generate antibodies
 CC which are used to purify, detect and target the polypeptides, including
 CC both in vivo and in vitro diagnostic and therapeutic methods. The
 CC proteins can be used to treat or prevent neural, immune system, muscular,
 CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
 CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
 CC sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 1931 BP; 496 A; 499 C; 517 G; 413 T; 6 other:
 XX
 XX Query Match 80.0%; Score 16.8; DB 21; Length 1931;
 XX Best Local Similarity 90.0%; Pred. No. 57;
 XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 2 gaagagacatgaacaactt 21
 ||||||| ||||||| ||
 Db 1658 gaagagacctggaacaaglt 1677
 RESULT 8
 AAF44684
 ID AAF44684 standard; cDNA; 2163 BP.
 AC AAF44684;
 XX

DT 27-MAR-2001 (first entry)

XX DE Novel protein kinase cDNA, SEQ ID NO: 64.

KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
 KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;
 KW immune disorder; cardiovascular disease; neurodegenerative disease;
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.

XX OS Homo sapiens.

XX PN WO200073469-A2.

XX PD 07-DEC-2000.

XX PF 26-MAY-2000; 2000WO-US14842.

XX PR 28-MAY-1999; 99US-0136503.

XX PA (SUGEN-) SUGEN INC.

XX PI Plowman GD, Martinez R, Whyte D, Sudersanam S;
 DR WPI; 2001-032161/04.
 DR P-PSDB; AAB65657.

XX PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers -

XX PS Disclosure: Fig 2; 310pp; English.

XX CC The present sequence encodes a novel protein kinase. The nucleic acids
 CC and the protein kinases they encode may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC and kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomypopathies, strokes, renal failure,
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 CC reproductive disorders.

XX SQ Sequence 2163 BP; 520 A; 572 C; 589 G; 482 T; 0 other;

Query Match 80.0%; Score 16.8; DB 22; Length 2163;
 Best Local Similarity 90.0%; Pred. No. 58;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 gaagagaccatgaacactt 21
 ||||| ||||| ||
 DB 1646 gaagagacctgaacagt 1665

RESULT 9
 ID AAV59121 standard; DNA; 2241 BP.
 AC AAV59121;
 XX 07-JAN-1999 (first entry)
 DT Nucleotide sequence of BCOM3, a kinase.
 DE

XX KW Serine protease; regulation; cell activity; viability; HELA2; ATC2;
 KW BCOM3; testisin; fertility; suppressor; testicular germ cell cancer;
 KW seminoma; testis-specific expression; antitumour; sperm development;
 KW infertility; ss.

XX OS Homo sapiens.

XX PN WO9836054-A1.

XX PD 20-AUG-1998.

XX PF 13-FEB-1998; 98WO-AU00085.

XX PR 18-NOV-1997; 97AU-0000422.
 PR 13-FEB-1997; 97AU-0005101.

XX PA (AMRA-) AMRAD OPERATIONS PTY LTD.
 XX PI Antalis TM, Hooper JD;
 DR WPI; 1998-480768/41.
 DR P-PSDB; AAM77299.

XX PT New serine protease(s) and kinase involved in regulating cell
 PT activity and viability - particularly the testis-specific protease
 PT HELA2 used for modulation of fertility and as tumour suppressor

XX PS Claim 30; Pages 70-73; 167pp; English.

XX CC The present sequence represents the nucleotide sequence of BCOM3.
 CC cDNA generated from HeLa cells and PAI-2 expressing HeLa cells
 CC was amplified using PCR primers AAV48312-13. Three new sequences
 CC were detected in the 480 bp amplicon. These sequences are designated
 CC HELA2 and ATC2 which have high homology to serine proteases and BCOM3
 CC which has homology to a kinase. The proteins are involved in or
 CC associated with regulation of cell activity and/or viability.
 CC Administration of recombinant HELA2 (also called testisin) is used to
 CC increase fertility. Downregulation of HELA2 reduces fertility. HELA2 is
 CC also a suppressor of testicular germ cell cancers (seminoma) and is also
 CC expressed in some non-testicular cancers (of colon, pancreas, prostate
 CC and ovary), so is a marker/potential therapeutic target for cancer. The
 CC promoter from the HELA2 gene is useful for testis-specific expression of
 CC other genes, e.g. for gene therapy or modulation of fertility. Drugs
 CC that block activity of HELA2 should have antitumour activity (other than
 CC in testis) while in testis recombinant HELA2 should stop growth of
 CC tumours and normalise sperm development (eliminating the need for
 CC orchidectomy). Identification of mutant forms of HELA2 can be used to
 CC diagnose infertility.

XX SQ Sequence 2241 BP; 557 A; 583 C; 602 G; 499 T; 0 other;

Query Match 80.0%; Score 16.8; DB 19; Length 2241;
 Best Local Similarity 90.0%; Pred. No. 58;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 gaagagaccatgaacactt 21
 ||||| ||||| ||
 DB 1699 gaagagacctgaacagt 1718

RESULT 10
 ID AAQ42188 standard; DNA; 1815 BP.
 AC AAQ42188;
 XX AAQ42188;
 XX

| | | |
|----|--|--|
| DT | 14-SEP-1993 | (first entry) |
| XX | ODC cDNA. | |
| DE | | |
| XX | Polymerase chain reaction; primer; amplify; PCR; protein A; rabbit; | |
| KW | ODC; hybridoma; ornithine decarboxylase; human; kidney; cDNA library; | |
| KW | fuslon protein; E. coli; antibody; detect; cancer; immunostaining; ss. | |
| XX | | |
| OS | Synthetic. | |
| XX | | |
| EH | Key | Location/Qualifiers |
| FT | CDS | 88..1473 |
| FT | | /*tag= a |
| PN | EP542287-A. | |
| PD | 19-MAY-1993. | |
| XX | | |
| PF | 13-NOV-1992; | 92EP-0119416. |
| XX | | |
| PR | 14-NOV-1991; | 91JP-0327041. |
| PR | 10-AUG-1992; | 92JP-0234278. |
| XX | | |
| PA | (IDEX) IDEMITSU KOSAN CO LTD. | |
| PL | Irie S, Nakayama K, Wadama K; | |
| DR | WPI; 1993-160785/20. | |
| DR | P-PSDB; AAR37270. | |
| XX | | |
| PT | Antibodies specific for human ornithine decarboxylase - useful | |
| PT | for diagnosis of cancer | |
| XX | | |
| PS | Disclosure; Fig 7; 23pp; English. | |
| XX | | |
| CC | This sequence represents the ornithine decarboxylase (ODC) cDNA. The | |
| CC | ODC cDNA was isolated using the primers given in AA042186-87. These | |
| CC | primers introduced a BamHI site to the 5' end, and an EcoRI site to | |
| CC | the 3' end of the ODC cDNA to facilitate cloning. A human kidney cDNA | |
| CC | library was used as the template and the ODC cDNA was isolated as a | |
| CC | BamHI-EcoRI fragment. This fragment was expressed in a protein A | |
| CC | fuslon gene vector. The vector was cloned in to E. coli and a protein | |
| CC | A-ODC fusion protein was produced. The ODC was sequenced, and | |
| CC | fusion protein and the N-terminus of the ODC was sequenced, and | |
| CC | compared to a reference N-terminal peptide to ensure that the protein | |
| CC | A had been cleaved. The isolated ODC protein was used to raise | |
| CC | antibodies in rabbits. The antibodies may be used to detect cancer | |
| CC | cells producing ODC by immunostaining. | |
| XX | | |
| SQ | Sequence 1815 BP; 485 A; 364 C; 449 G; 517 T; 0 other; | |
| | | |
| | Query Match | 75.2%; Score 15.8; DB 14; Length 1815; |
| | Best Local Similarity | 89.5%; Pred. NO. 1.7e+02; |
| | Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0; | |
| Oy | 3 aagagacatgacaactt 21 | |
| Db | 80 aagaatacatgacaactt 98 | |
| | | |
| | RESULT 11 | |
| ID | AAF18267 | AAF18267 |
| XX | AAF18267 standard; DNA; 1958 BP. | |
| XX | | |
| AC | AAF18267; | |
| XX | | |
| DT | 14-MAR-2001 (first entry) | |
| XX | | |
| DE | Lung cancer associated, polynucleotide sequence SEQ ID 286. | |
| XX | | |
| KW | Human; lung cancer associated protein; neuroprotective; cytostatic; | |
| KW | cardioactive; immunomodulatory; muscular active; vulnerary; | |

| | |
|-----------|--|
| KW | gastrointestinal; nephrotropic; antineoplastic; gynecological; |
| KW | antibacterial; diagnosis; neutral disorder; immune disorder; reproductive; |
| KW | proliferative disorder; wound healing; infectious disease; ds. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | MO200055180-A2. |
| PD | 21-SEP-2000. |
| XX | |
| PE | 08-MAR-2000; 2000MO-US05918. |
| XX | |
| PR | 12-MAR-1999; 99US-0124270. |
| XX | |
| PA | (HUMA-) HUMAN GENOME SCI INC. |
| XX | |
| PA | (ROSE/) ROSEN C. A. |
| XX | |
| PI | Ruben SM; |
| XX | |
| DR | WPI; 2000-587514/55. |
| XX | |
| DR | P-PSDB; AAB58391. |
| XX | |
| PT | Lung cancer associated gene sequences, referred to as lung cancer |
| XX | |
| PT | antigens, useful for treatment, prevention, and diagnosis of disorders |
| XX | |
| PT | such as lung cancer - |
| XX | |
| PS | Claim 1; Page 744-745; 1425pp; English. |
| XX | |
| CC | Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer |
| XX | |
| CC | associated proteins represented in AAB50106 - AAB58548. Lung cancer |
| XX | |
| CC | associated proteins and polynucleotide sequences, their agonists, and |
| XX | |
| CC | antagonists may have neuroprotective; cytoskeletal; cardiovascular; |
| XX | |
| CC | immunomodulatory; muscular active general; vulnery; gastrointestinal |
| XX | |
| CC | general; nephrotropic; antineoplastic; gynecological; or antibacterial |
| XX | |
| CC | activity. The invention also includes antibodies specific for the |
| XX | |
| CC | protein or polynucleotide sequences. The lung cancer associated |
| XX | |
| CC | polynucleotide sequences may be used for detection of lung cancer, |
| XX | |
| CC | chromosome identification, as chromosome markers, and for numerous other |
| XX | |
| CC | diagnostic or research purposes. The proteins may be used to treat |
| XX | |
| CC | disorders such as neural, immune, muscular, reproductive, |
| XX | |
| CC | gastrointestinal, pulmonary, cardiovascular, renal, and proliferative |
| XX | |
| CC | disorders. The proteins may also be used in the treatment of wounds and |
| XX | |
| CC | infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and |
| XX | |
| CC | peptide AAB5549 are used in the course of the invention for the |
| XX | |
| CC | identification and characterisation of the polynucleotide and protein |
| XX | |
| CC | sequences. |
| XX | |
| SQ | Sequence 1958 BP; 520 A; 402 C; 497 G; 538 T; 1 other; |
| XX | |
| QY | Query Match 75.2%; Score 15.8; DB 21; Length 1958; |
| XX | |
| QY | Best Local Similarity 89.5%; Pred. No. 1.7e+02; |
| XX | |
| QY | Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0; |
| XX | |
| QY | 3 aagagaccatgaacaactt 21 |
| XX | |
| DB | 171 aagaatcatgaacaactt 189 |
| XX | |
| RESULT 12 | |
| AAZ44288 | |
| ID | AAZ44288 standard; cDNA; 2265 BP. |
| XX | |
| AC | AAZ44288; |
| XX | |
| DT | 31-MAR-2000 (first entry) |
| XX | |
| DE | Human brain PARP3 cDNA. |
| XX | |
| KW | PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis; |
| XX | |
| KW | diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness; |
| XX | |
| KW | ischemic tissue damage; PARP3; ss. |
| XX | |

OS Homo sapiens.
 XX Key Location/Qualifiers
 FH CDS 242..1843
 FT /*tag= a
 FT /product= "PARP3"
 FT W09964572-A2.
 XX 16-DEC-1999.
 PD 04-JUN-1999; 99WO-EP03889.
 XX 05-JUN-1998; 98DE-1025213.
 PR 01-MAR-1999; 99DE-1008837.
 XX (BADI) BASF AG.
 PA Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
 PI WPI: 2000-087218/07.
 DR P-PSDB: AAY51175.
 XX Novel genes and proteins, antibodies and binding partners useful in
 PT diagnosis and therapy of energy deficiency associated disease
 PT conditions -
 PS Claim 7b; Page 54-57; 96pp; German.
 XX This invention describes novel human and murine poly(ADP-ribose)
 CC polymerase (PARP) homologues, which are characterised by an amino acid
 CC sequence with a functional NAD⁺-binding site and no zinc finger
 CC sequence motif, of general formula CX₂CX₂MHX₂C (1). The nucleic acid
 CC sequences, PARP homologues and antibodies are useful for analytic
 CC detection of PARP homologues and for identifying PARP effectors or
 CC binding partners, as well as for determining their effectiveness.
 CC PARP-binding partners are useful for the diagnosis or therapy of a
 CC disease condition, which is the result of a PARP protein, especially an
 CC energy deficiency, which may comprise tissue damage from cell death
 CC following necrosis or apoptosis. The disease condition may be chosen
 CC from a neurodegenerative illness, or sepsis or ischemic tissue damage,
 CC in particular neurotoxic disturbances, etc. This sequence encodes the
 CC human PARP3 protein used in the method of the invention.
 SQ Sequence 2265 BP; 570 A; 662 C; 587 G; 446 T; 0 other;
 QY 2 gaagagaccatgaacaact 20
 DB 632 gaagaagaccagaacaact 650
 RESULT 13
 ID AA244289 standard; CDNA; 2265 BP.
 XX AA244289;
 AC 31-MAR-2000 (first entry)
 DE Human uterus type 2 PARP3 CDNA.
 XX PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
 KW diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
 KW ischemic tissue damage; PARP3; ss.
 XX Homo sapiens.
 OS Homo sapiens.
 FH Key Location/Qualifiers

FT CDS 221..1843
 FT /*tag= a
 FT /product= "PARP3"
 FT W09964572-A2.
 XX 16-DEC-1999.
 PD 04-JUN-1999; 99WO-EP03889.
 XX 05-JUN-1998; 98DE-1025213.
 PR 01-MAR-1999; 99DE-1008837.
 XX (BADI) BASF AG.
 PA Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
 PI WPI: 2000-087218/07.
 DR P-PSDB: AAY51176.
 XX Novel genes and proteins, antibodies and binding partners useful in
 PT diagnosis and therapy of energy deficiency associated disease
 PT conditions -
 PS Claim 7c; Page 59-62; 96pp; German.
 XX This invention describes novel human and murine poly(ADP-ribose)
 CC polymerase (PARP) homologues, which are characterised by an amino acid
 CC sequence with a functional NAD⁺-binding site and no zinc finger
 CC sequence motif, of general formula CX₂CX₂MHX₂C (1). The nucleic acid
 CC sequences, PARP homologues and antibodies are useful for analytic
 CC detection of PARP homologues and for identifying PARP effectors or
 CC binding partners, as well as for determining their effectiveness.
 CC PARP-binding partners are useful for the diagnosis or therapy of a
 CC disease condition, which is the result of a PARP protein, especially an
 CC energy deficiency, which may comprise tissue damage from cell death
 CC following necrosis or apoptosis. The disease condition may be chosen
 CC from a neurodegenerative illness, or sepsis or ischemic tissue damage,
 CC in particular neurotoxic disturbances, etc. This sequence encodes the
 CC human PARP3 protein used in the method of the invention.
 SQ Sequence 2265 BP; 570 A; 662 C; 587 G; 446 T; 0 other;
 QY 2 gaagagaccatgaacaact 20
 DB 632 gaagaagaccagaacaact 650
 RESULT 14
 ID AAC63441/C
 XX AAC63441 standard; CDNA; 1913 BP.
 AC AAC63441;
 DE 09-FEB-2001 (first entry)
 DE Human secreted protein coding sequence SEQ ID NO: 42.
 XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
 KW vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; ss.
 XX Homo sapiens.
 OS Homo sapiens.
 FH WO200061779-A1.

```

PD 19-OCT-2000.
XX
XX 06-APR-2000; 2000MO-US09068.
XX
XX 09-APR-1999; 99US-0128699.
XX
XX 20-JAN-2000; 2000US-0177050.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsoulis G;
XX
XX WPI; 2000-647424/62.
XX
XX P-PSDB; AAB29833.
XX
XX Isolated nucleic acid molecule encoding a human secreted protein is
XX PT used in preventing, treating or ameliorating a medical condition
XX
XX
XX PS Claim 1; Page 414; 495pp; English.
XX
XX The invention relates to the isolation of genes AAGC3410-CG3458 encoding
XX CC 49 human secreted proteins AAB29802-B29850. The genes can be used to
XX CC generate fusion proteins by linking to the gene for the human
XX CC immunoglobulin G Fc portion (SEQID) for increasing the stability of
XX CC the fusion protein as compared to the human protein only. The genes and
XX CC proteins are useful for preventing, ameliorating or treating medical
XX CC conditions, e.g. by protein or gene therapy. The genes are isolated
XX CC from a range of human tissues disclosed in the specification. The
XX CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX CC and ovarian cancer, and other cancers of the adrenal gland, bone, bone
XX CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
XX CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX CC colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
XX CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
XX CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
XX CC and parasitic infections.
XX
XX SQ Sequence 1913 BP; 601 A; 356 C; 375 G; 581 T; 0 other;

Query Match          73.3%; Score 15.4; DB 21; Length 1913;
Best Local Similarity 94.1%; Pred. No. 2.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 gagagacacatgaacaa 18
   |||||
Db 1112 GAAGAGACATGCACAA 1096

RESULT 15
AAQ04775/c
ID AAQ04775 standard; DNA; 1915 BP.
XX
XX AAQ04775:
AC 18-OCT-1990 (first entry)
XX
XX 18-OCT-1990 (first entry)
XX
XX Fragment contg. Chloramphenicol resistance gene.
DE
XX
XX Chloramphenicol resistance; mutagenesis; ss.
KM
XX
XX Key Location/Qualifiers
FH CDS 545..1714
FT /*tag= a
FT /label=Chloramphenicol resistance gene
FT RBS 535..538
FT /*tag= b
XX
XX DE3841454-A.
XX
XX 13-JUN-1990.

```

```

XX
XX 09-DEC-1988; 88DE-3841454.
XX
XX 09-DEC-1988; 88DE-3841454.
XX
XX (DEGS) DEGUSSA AG.
XX
XX Kassing F, Kalinowski J, Arnold W, Winterfeldt A, Puhler A;
XX PI Kautz P-S, Thierbach G;
XX
XX WPI; 1990-186465/25.
XX
XX P-PSDB; AAR05475.
XX
XX Site specific mutagenesis of DNA at restriction enzyme positions -
XX PT by treating corresp. restriction fragments with hydroxylamine,
XX PT and new mutated sequences and transformed microorganisms
XX
XX
XX PS Disclosure; ; P; German.
XX
XX This fragment when ligated into a plasmid and used to transform host
XX CC cells confers Chloramphenicol resistance on them.
XX
XX SQ Sequence 1915 BP; 353 A; 601 C; 567 G; 394 T; 0 other;

Query Match          73.3%; Score 15.4; DB 11; Length 1915;
Best Local Similarity 94.1%; Pred. No. 2.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 gagacacatgaacact 21
   |||||
Db 1826 GAGACATGAACATCTT 1810

Search completed: September 21, 2001, 02:42:51
Job time: 11467 sec

```

U

U

U

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2001, 23:28:33 ; Search time 4309.39 Seconds
(without alignments)
46.065 Million cell updates/sec

Title: US-09-138-735-12

Perfect score: 21
Sequence: 1 cgaagagaccatgaacaactt 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:**

1: gb_est1:**
2: gb_est2:**
3: gb_est3:**
4: gb_est4:**
5: gb_est5:**
6: gb_est6:**
7: gb_est7:**
8: gb_est8:**
9: gb_est9:**
10: gb_est10:**
11: gb_est11:**
12: gb_est12:**
13: gb_est13:**
14: gb_est14:**
15: gb_est15:**
16: gb_est16:**
17: gb_est17:**
18: gb_est18:**
19: gb_est19:**
20: gb_est20:**
21: gb_est21:**
22: gb_est22:**
23: gb_est23:**
24: gb_est24:**
25: gb_est25:**
26: gb_est26:**
27: gb_est27:**
28: gb_est28:**
29: gb_est29:**
30: gb_est30:**
31: gb_est31:**
32: gb_est32:**
33: gb_est33:**
34: gb_est34:**
35: gb_est35:**
36: gb_est36:**
37: gb_est37:**
38: gb_est38:**
39: gb_est39:**
40: gb_est40:**
41: gb_est41:**
42: gb_est42:**
43: gb_est43:**
44: gb_est44:**
45: gb_est45:**
46: gb_est46:**
47: gb_est47:**

44: em_esthum10:**
45: em_esthum11:**
46: em_esthum12:**
47: em_esthum13:**
48: em_esthum14:**
49: em_esthum15:**
50: em_esthum16:**
51: em_esthum17:**
52: em_esthum18:**
53: em_esthum19:**
54: em_esthum20:**
55: em_esthum21:**
56: em_esthum22:**
57: em_esthum23:**
58: em_esthum24:**
59: em_esthum25:**
60: em_esthum26:**
61: em_esthum27:**
62: em_esthum28:**
63: em_esthum29:**
64: em_esthum30:**
65: em_esthum31:**
66: em_esthum32:**
67: em_esthum33:**
68: em_esthum34:**
69: em_esthum35:**
70: em_esthum36:**
71: em_esthum37:**
72: em_esthum38:**
73: em_esthum39:**
74: em_esthum40:**
75: em_esthum41:**
76: em_esthum42:**
77: em_esthum43:**
78: em_esthum44:**
79: em_esthum45:**
80: em_esthum46:**
81: em_esthum47:**
82: em_esthum48:**
83: em_esthum49:**
84: em_esthum50:**
85: em_esthum51:**
86: em_esthum52:**
87: em_esthum53:**
88: em_esthum54:**
89: em_esthum55:**
90: em_esthum56:**
91: em_esthum57:**
92: em_esthum58:**
93: em_esthum59:**
94: em_esthum60:**
95: em_esthum61:**
96: em_esthum62:**
97: em_esthum63:**
98: em_esthum64:**
99: em_esthum65:**
100: em_esthum66:**
101: em_esthum67:**
102: em_esthum68:**
103: em_esthum69:**
104: em_esthum70:**
105: em_esthum71:**
106: em_esthum72:**
107: em_esthum73:**
108: em_esthum74:**
109: em_esthum75:**
110: em_esthum76:**
111: em_esthum77:**
112: em_esthum78:**
113: em_esthum79:**
114: em_esthum80:**
115: em_esthum81:**
116: em_esthum82:**

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 18.4 | 87.6 | 504 | 14 AA982948 | AA982948 ua22h06.r |
| 2 | 18.4 | 87.6 | 520 | 118 AM585823 | EST317446 |
| 3 | 18.4 | 87.6 | 533 | 171 BF932609 | BF932609 RC5-PT018 |
| 4 | 17.8 | 84.8 | 532 | 139 BE760064 | BE760064 an-2069 A |
| 5 | 17.4 | 82.9 | 628 | 141 BE856251 | BE856251 7191c06.x |
| 6 | 16.8 | 80.0 | 118 | 120 AM805050 | AM805050 QV1-UM009 |
| 7 | 16.8 | 80.0 | 209 | 169 BF799621 | BF799621 CM0-C1009 |
| 8 | 16.8 | 80.0 | 277 | 1 AA043840 | AA043840 zk6ia12.r |
| 9 | 16.8 | 80.0 | 282 | 168 BF738704 | BF738704 PM4-KT004 |
| 10 | 16.8 | 80.0 | 298 | 159 N78192 | N78192 yv74e05.r1 |
| 11 | 16.8 | 80.0 | 300 | 26 AV177260 | AV177260 AV177260 |
| 12 | 16.8 | 80.0 | 300 | 26 AV178859 | AV178859 AV178859 |
| 13 | 16.8 | 80.0 | 300 | 26 AV184752 | AV184752 AV184752 |
| 14 | 16.8 | 80.0 | 300 | 156 C29371 | C29371 C29371 Yuj1 |
| 15 | 16.8 | 80.0 | 300 | 156 C29371 | C29371 C29371 Yuj1 |
| 16 | 16.8 | 80.0 | 300 | 156 C37295 | C37295 C37295 Yuj1 |
| 17 | 16.8 | 80.0 | 300 | 156 C37468 | C37468 C37468 Yuj1 |
| 18 | 16.8 | 80.0 | 300 | 156 C38177 | C38177 C38177 Yuj1 |
| 19 | 16.8 | 80.0 | 300 | 156 C38520 | C38520 C38520 Yuj1 |
| 20 | 16.8 | 80.0 | 309 | 170 BF856629 | BF856629 PM1-PT020 |
| 21 | 16.8 | 80.0 | 323 | 170 BF857651 | BF857651 PM1-PT020 |
| 22 | 16.8 | 80.0 | 325 | 170 BF856627 | BF856627 PM1-PT020 |
| 23 | 16.8 | 80.0 | 331 | 170 BF85438 | BF85438 RC6-TN007 |
| 24 | 16.8 | 80.0 | 347 | 139 BE693578 | BE693578 RC3-PT033 |
| 25 | 16.8 | 80.0 | 347 | 170 BF856628 | BF856628 PM1-PT020 |
| 26 | 16.8 | 80.0 | 354 | 170 BF856609 | BF856609 PM1-PT020 |
| 27 | 16.8 | 80.0 | 357 | 170 BF856632 | BF856632 PM1-PT020 |
| 28 | 16.8 | 80.0 | 359 | 170 BF857646 | BF857646 PM1-PT020 |
| 29 | 16.8 | 80.0 | 360 | 157 D67597 | D67597 CELK034E3F |
| 30 | 16.8 | 80.0 | 361 | 170 BF856634 | BF856634 PM1-PT020 |
| 31 | 16.8 | 80.0 | 367 | 170 BF856611 | BF856611 PM1-PT020 |
| 32 | 16.8 | 80.0 | 368 | 115 AM380268 | AM380268 QV3-HT026 |
| 33 | 16.8 | 80.0 | 372 | 140 BE817820 | BE817820 CM0-BN026 |
| 34 | 16.8 | 80.0 | 373 | 115 AM380277 | AM380277 QV3-HT026 |
| 35 | 16.8 | 80.0 | 378 | 122 AM917829 | AM917829 EST349133 |
| 36 | 16.8 | 80.0 | 388 | 170 BF857647 | BF857647 PM1-PT020 |
| 37 | 16.8 | 80.0 | 400 | 140 BE830145 | BE830145 RC6-FT008 |
| 38 | 16.8 | 80.0 | 422 | 9 AA587314 | AA587314 nnt4q11.s |
| 39 | 16.8 | 80.0 | 433 | 14 AA975653 | AA975653 c063e02.s |
| 40 | 16.8 | 80.0 | 434 | 116 AM482284 | AM482284 41788 MAR |
| 41 | 16.8 | 80.0 | 444 | 158 H79060 | H79060 yu24e03.r1 |
| 42 | 16.8 | 80.0 | 475 | 189 W01185 | W01185 za59e08.r1 |
| 43 | 16.8 | 80.0 | 484 | 123 AM984006 | AM984006 RC1-HN000 |
| 44 | 16.8 | 80.0 | 495 | 3 AA196613 | AA196613 zq08a03.r |
| 45 | 16.8 | 80.0 | 495 | 123 AM983992 | AM983992 RC1-HN000 |

ALIGNMENTS

RESULT 1
LOCUS AA982948 504 bp mRNA EST 27-MAY-1998
DEFINITION ua22h06.r1 Soares_mammmary_gland_NBMWG Mus musculus cDNA clone
IMAGE:1347515 5' similar to gb:X92346 M.musculus mRNA for CART1
protein (MOUSE);, mRNA sequence.
ACCESSION AA982948
VERSION AA982948
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 504)
Matta,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

TITLE
JOURNAL
COMMENT

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The Mashu-HIMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
Mashu-HIMI Mouse EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:696307
Trace considered overall poor quality
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..504
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_image="1347515"
/clone_lib="Soares_mammmary_gland_NBMWG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dt)
primer [5',
TGTACCACTGAGTGGAGCGCCGCAAGCTTTTGTGTGTGTGTGTGTGTGT
T3); double stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

BASE COUNT
ORIGIN

129 a 129 c 131 g 115 t

Query Match 87.6%; Score 18.4; DB 14; Length 504;
Best Local Similarity 95.0%; Pred. No. 1.3e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 cgaagagacatgaact 20
1 |||
Db 172 CTAGAGACATGACACT 191

RESULT 2

AM585823 520 bp mRNA EST 07-SEP-2000
DEFINITION EST317446 MHAM Medicago truncatula/Gloms versiforme mixed EST
library cDNA clone PMHAM-39L1, mRNA sequence.
ACCESSION AM585823
VERSION AM585823
KEYWORDS EST.
SOURCE Medicago truncatula/Gloms versiforme mixed EST library
ORGANISM Eukaryota; mixed EST libraries.
REFERENCE 1 (bases 1 to 520)
Harrison,M.J., Liu,J., Peng,H., Gonzales,M., Ellis,L., Town,C.D.,
Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and Fraser,C.M.
ESTs from roots of Medicago truncatula after colonization with
Gloms versiforme
Unpublished (2000)
Contact: Maria J. Harrison
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73401, USA

Tel: 580-223-5810
 Fax: 580-221-7380
 Email: mjharrison@noble.org
 Noble EST name: N254751e
 TIGR sequence name: MTDAG617K
 More information is available at:
 'http://chryslie.tamu.edu/medicago/
 Seq primer: SKmod (CTA gAA CTA gTg gAT CC).
 Location/Qualifiers

FEATURES

source

1. 520
 /organism="Medicago truncatula/Gloms versiforme mixed EST
 library"
 /cultivar="Medicago truncatula genotype A17"
 /db_xref="taxon:119092"
 /clone_1lb="PMHAM-39L1"
 /clone_1lb="MHAM"
 /tissue_type="roots colonized with Gloms versiforme"
 /dev_stage="roots harvested at 10, 17, 22, 31 and 38 days
 post-inoculation with Gloms versiforme. The library was
 made from a mixture of RNA from each of these stages."
 /lab_host="E. coli strain XL0R"
 /note="Vector: Bluescript SK-; Site 1: EcoRI; Site 2:
 XhoI; cDNA was prepared from polyA+ enriched RNA from
 roots harvested at 10, 17, 22, 31 and 38 days
 post-inoculation with Gloms versiforme. The cDNA was
 directionally ligated into the Unizap XR vector from
 StrataGene and packaged using Gigapack III Gold packaging
 extracts. Plasmids containing cDNA inserts were excised
 from the recombinant lambda-Zap phage using Ex-assist
 helper phage and propagated in XL0R cells."

BASE COUNT
 ORIGIN

178 a 96 c 95 g 151 t

Query Match

Best Local Similarity 87.6%; Score 18.4; DB 118; Length 520;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 gaagagacatgaacactt 21
 |||||

Db 48 GAAGAGACCTGAACACTT 67

RESULT 3
 LOCUS BE932609 533 bp mRNA EST 22-JAN-2001
 DEFINITION RC5-WT0181-201200-021-F10 WT0181 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE932609
 VERSION BE932609.1 GI:12349933
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 533)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC5&e2=RC5-WT0181-
 201200-021-F10&tl3=2000-12-20&tl4=1)
 Seq primer: puc 18 forward
 high quality sequence start: 43
 high quality sequence stop: 384.
 Location/Qualifiers

FEATURES

source

1. 533
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1lb="WT0181"
 /dev_stage="Adult"
 /note="Organ: nervous_tumor; Vector: puc18; site_1: SmaI;
 site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent Application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

BASE COUNT
 ORIGIN

135 a 130 c 163 g 105 t

Query Match

Best Local Similarity 87.6%; Score 18.4; DB 171; Length 533;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 gaagagacatgaacactt 21
 |||||

Db 322 GAAGAGACCTGAACACTT 303

RESULT 4
 LOCUS BE760064 532 bp mRNA EST 24-OCT-2000
 DEFINITION an_2069 Aspergillus niger, pYES2 (XhoI-EcoRI) Aspergillus niger
 cDNA clone 2069 5', mRNA sequence.
 ACCESSION BE760064
 VERSION BE760064.1 GI:10182701
 KEYWORDS EST.
 SOURCE Aspergillus niger.
 ORGANISM Aspergillus niger.
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 1 (bases 1 to 532)
 Tsang, A. and Storms, R.
 Aspergillus niger Expressed Sequence Tags
 Unpublished (2000)
 CONTACT: Tsang, A.
 Dept. of Biology
 Concordia University
 1455 de Maisonneuve Blvd. West, Montreal, Quebec CANADA, H3G 1M8
 Tel: 514-848-3405
 Fax: 514-848-4504
 Email: tsang@vax2.concordia.ca
 PCR PRIMERS
 FORWARD: CGACATTATGCTGAGTGAATATCC
 Insert Length: 532 Std Error: 0.00
 POLYA=No.

FEATURES

source

1. 532
 /organism="Aspergillus niger"
 /db_xref="taxon:5061"
 /clone="2069"
 /clone_1lb="Aspergillus niger, pYES2 (XhoI-EcoRI)"
 /lab_host="E. coli"
 /note="Vector: pYES2; Site 1: Xho-I; Site 2: EcoRI; cDNA
 was synthesized with Zap kit (Stratagene) using poly(A)+
 RNA isolated from Aspergillus niger grown in both complete
 and minimal media. Synthesis was primed with oligo(dT)
 primer/XhoI-linker. EcoRI adaptors were later ligated to
 polished ends. EcoRI-XhoI-digested cDNA was ligated with

SOURCE human.
ORGANISM Homo sapiens
COMMENT Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 209)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovsky-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=CM0&t2=CM0-CI0096-101000-599-a10&t3=2000-10-10&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 209.
Location/Qualifiers
FEATURES
source
1..209
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CI0096"
/dev_stage="Adult"
/note="Organ: colon.ins; Vector: puc18; Site.1: Sma1; Site.2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 58 a 45 c 67 g 39 t
ORIGIN
Query Match 80.0%; Score 16.8; DB 169; Length 209;
Best Local Similarity 90.0%; Pred. No. 6.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 gaagagacatgaacaact 21
|||||
Db 68 GAAGAGACCTTGAAACAGTT 49
RESULT 8
LOCUS AA043840 277 bp mRNA EST 04-SEP-1996
DEFINITION zK61a12.r1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone
IMAGE:487294 5', mRNA sequence.
ACCESSION AA043840
VERSION AA043840.1 GI:1521697
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
COMMENT Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 277)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 247.
Location/Qualifiers
FEATURES
source
1..277
/organism="Homo sapiens"
/db_xref="GDB:376108"
/db_xref="taxon:9606"
/clone_lib="IMAGE:487294"
/clone_lib="Soares_pregnant_uterus_NbHPU"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pRTT3-Pac; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' ACTGTGAAGATTCGCGCCGCTTTTCTTTTCTTTT 3']
double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRTT3 vector. Library went through one round of normalization. Library constructed by M. Palma Bonaldo."
BASE COUNT 73 a 63 c 83 g 58 t
ORIGIN
Query Match 80.0%; Score 16.8; DB 1; Length 277;
Best Local Similarity 90.0%; Pred. No. 7.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 gaagagacatgaacaact 21
|||||
Db 228 GAAGAGACCTTGAAACAGTT 247
RESULT 9
LOCUS BF738704 282 bp mRNA EST 10-JAN-2001
DEFINITION PM4-KT0045-191200-005-a07 KT0045 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF738704
VERSION BF738704.1 GI:12065380
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
COMMENT Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 282)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovsky-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

RESULT 12
 AV178859 300 bp mRNA EST 21-JUL-1999
 LOCUS AV178859 Yui1 Kohara unpublished cDNA:Strain N2 hermaphrodite
 DEFINITION embryo Caenorhabditis elegans cDNA clone yk573a1 3', mRNA sequence.
 ACCESSION AV178859 GI:5556760
 VERSION AV178859.1 GI:5556760
 SOURCE EST.
 KEYWORDS Caenorhabditis elegans.
 ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea
 ; Rhabdilitidae; Pelodierinae; Caenorhabditis.
 REFERENCE 1 (bases 1 to 300)
 AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Mitsuiki,H.,
 Nishigaki,A., Motohashi,T., Zeng,Q., Watanabe,H., Sugimoto,A., Sano
 ,M., Miyata,A., Mitani,Y., Iida,K., Uesugi,H., Sugiyama,Y. and
 Nomoto,H.
 TITLE Expressed genes in C.elegans
 JOURNAL Unpublished (1999)
 COMMENT Contact: Yui1 Kohara
 Gene Library Lab
 National Institute of Genetics
 Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 81-559-81-6854
 Fax: 81-559-81-6855
 Email: ykohara@lab.nig.ac.jp.
 Location/Qualifiers
 1..300
 /organism="Caenorhabditis elegans"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="yk573a1"
 /clone_1lb="Yui1 Kohara unpublished cDNA:Strain N2
 hermaphrodite embryo"
 /sex="hermaphrodite"
 /dev_stage="embryo"
 /dev_stage="embryo"

BASE COUNT 107 a 56 c 65 g 67 t 5 others
 ORIGIN

Query Match 80.0%; Score 16.8; DB 26; Length 300;
 Best Local Similarity 90.0%; Pred. No. 7.3e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 gaagagaccatgaacaactt 21
 ||||| ||||| ||||| |||||
 Db 233 GAAGAAACATGCACACTT 252

RESULT 13
 AV184752 300 bp mRNA EST 21-JUL-1999
 LOCUS AV184752 Yui1 Kohara unpublished cDNA:Strain N2 hermaphrodite
 DEFINITION embryo Caenorhabditis elegans cDNA clone yk669a12 3', mRNA
 sequence.
 ACCESSION AV184752 GI:5564653
 VERSION AV184752.1 GI:5564653
 SOURCE EST.
 KEYWORDS Caenorhabditis elegans.
 ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea
 ; Rhabdilitidae; Pelodierinae; Caenorhabditis.
 REFERENCE 1 (bases 1 to 300)
 AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Mitsuiki,H.,
 Nishigaki,A., Motohashi,T., Zeng,Q., Watanabe,H., Sugimoto,A., Sano
 ,M., Miyata,A., Mitani,Y., Iida,K., Uesugi,H., Sugiyama,Y. and
 Nomoto,H.
 TITLE Expressed genes in C.elegans
 JOURNAL Unpublished (1999)
 COMMENT Contact: Yui1 Kohara
 Gene Library Lab
 National Institute of Genetics

Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 81-559-81-6854
 Fax: 81-559-81-6855
 Email: ykohara@lab.nig.ac.jp.
 Location/Qualifiers
 1..300
 /organism="Caenorhabditis elegans"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="yk669a12"
 /clone_1lb="Yui1 Kohara unpublished cDNA:Strain N2
 hermaphrodite embryo"
 /sex="hermaphrodite"
 /dev_stage="embryo"
 /dev_stage="embryo"

BASE COUNT 105 a 48 c 76 g 67 t 4 others
 ORIGIN

Query Match 80.0%; Score 16.8; DB 26; Length 300;
 Best Local Similarity 90.0%; Pred. No. 7.3e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 gaagagaccatgaacaactt 21
 ||||| ||||| ||||| |||||
 Db 278 GAAGAAACATGCACACTT 297

RESULT 14
 C29371 300 bp mRNA EST 18-OCT-1999
 LOCUS C29371 Yui1 Kohara unpublished cDNA:Strain N2 hermaphrodite embryo
 DEFINITION Caenorhabditis elegans cDNA clone yk201f1 3', mRNA sequence.
 ACCESSION C29371 GI:2361167
 VERSION C29371.1 GI:2361167
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans.
 ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea
 ; Rhabdilitidae; Pelodierinae; Caenorhabditis.
 REFERENCE 1 (bases 1 to 300)
 AUTHORS Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano
 ,M., Miyata,A. and Nishigaki,A.
 TITLE Expression map of the C.elegans genome
 JOURNAL Unpublished (1996)
 COMMENT Contact: Yui1 Kohara
 Gene Library Lab
 National Institute of Genetics
 Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 81-559-81-6854
 Fax: 81-559-81-6855
 Email: ykohara@lab.nig.ac.jp.
 Location/Qualifiers
 1..300
 /organism="Caenorhabditis elegans"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="yk201f1"
 /clone_1lb="Yui1 Kohara unpublished cDNA:Strain N2
 hermaphrodite embryo"
 /sex="hermaphrodite"
 /dev_stage="embryo"
 /dev_stage="embryo"

BASE COUNT 110 a 48 c 75 g 64 t 3 others
 ORIGIN

Query Match 80.0%; Score 16.8; DB 156; Length 300;
 Best Local Similarity 90.0%; Pred. No. 7.3e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 gaagagaccatgaacaactt 21
 ||||| ||||| ||||| |||||
 Db 279 GAAGAAACATGCACACTT 298

RESULT 15

LOCUS

C37295

300 bp mRNA

EST 18-OCT-1999

DEFINITION

C37295 Yuji Kohara unpublished cDNA:Strain N2 hermaphrodite embryo
Caenorhabditis elegans cDNA clone yk464a3 3', mRNA sequence.

ACCESSION

C37295

VERSION

C37295.1 GI:2373532

KEYWORDS

EST

SOURCE

Caenorhabditis elegans.

REFERENCE

AUTHORS

Kohara,Y., Motobashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano
,M., Miyata,A. and Nishigaki,A.
Expression map of the C.elegans genome
Unpublished (1996)

TITLE

JOURNAL

COMMENT

Gene Library Lab
National Institute of Genetics
Yata 1111 Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES

Location/Qualifiers

SOURCE

1..300
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk464a3"
/clone_id="Yuji Kohara unpublished cDNA:Strain N2
hermaphrodite embryo"
/sex="hermaphrodite"
/dev_stage="embryo"

| | | | | | |
|------------|-------|------|------|------|----------|
| BASE COUNT | 107 a | 50 c | 74 g | 66 t | 3 others |
| ORIGIN | | | | | |

Query Match

Best Local Similarity 80.0%; Score 16.8; DB 156; Length 300;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

| | | | |
|----|-----|----------------------|-----|
| Qy | 2 | gaagagaccatgaacacact | 21 |
| | | | |
| Db | 264 | GAAGAACCATGCACACACTT | 283 |

Search completed: September 20, 2001, 23:28:42
Job time: 25582 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 02:16:42 : Search time 164.23 Seconds
(without alignments)
24.207 Million cell updates/sec

Title: US-09-138-735-12

Perfect score: 21

Sequence: 1 cgaagagacatgacaaactt 21

Scoring table: IDENTITY_NUC

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn1_7/ptodata/1/ina/5A.COMB.seq: *
2: /cgn1_7/ptodata/1/ina/5B.COMB.seq: *
3: /cgn1_7/ptodata/1/ina/6A.COMB.seq: *
4: /cgn1_7/ptodata/1/ina/6B.COMB.seq: *
5: /cgn1_7/ptodata/1/ina/6C.COMB.seq: *
6: /cgn1_7/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------------|--------------------|
| 1 | 21 | 100.0 | 21 | US-08-480-917-12 | Sequence 12, Appl |
| 2 | 21 | 100.0 | 3402 | US-08-480-917-1 | Sequence 1, Appl |
| 3 | 16.8 | 80.0 | 2029 | US-08-933-750C-69 | Sequence 69, Appl |
| 4 | 16.8 | 80.0 | 2029 | US-09-234-613-69 | Sequence 69, Appl |
| 5 | 15.4 | 73.3 | 2533 | US-08-861-745B-2 | Sequence 2, Appl |
| 6 | 15.2 | 72.4 | 1046 | US-08-361-467B-4 | Sequence 4, Appl |
| 7 | 15.2 | 72.4 | 1046 | US-08-484-332C-4 | Sequence 4, Appl |
| 8 | 15.2 | 72.4 | 1610 | US-09-276-531-56 | Sequence 56, Appl |
| 9 | 15.2 | 72.4 | 1685 | US-08-360-673-1 | Sequence 1, Appl |
| 10 | 15.2 | 72.4 | 12492 | 5206163-2 | Patent No. 5206163 |
| 11 | 15 | 71.4 | 8906 | US-08-826-267-1 | Sequence 1, Appl |
| 12 | 14.8 | 70.5 | 69 | PCT-US92-01015-12 | Sequence 12, Appl |
| 13 | 14.8 | 70.5 | 72 | PCT-US92-01015-3 | Sequence 3, Appl |
| 14 | 14.8 | 70.5 | 81 | US-08-026-145-3 | Sequence 3, Appl |
| 15 | 14.8 | 70.5 | 96 | US-07-872-673B-4 | Sequence 4, Appl |
| 16 | 14.8 | 70.5 | 1545 | US-07-872-673B-2 | Sequence 2, Appl |
| 17 | 14.8 | 70.5 | 6216 | US-09-213-053-1 | Sequence 1, Appl |
| 18 | 14.8 | 70.5 | 7742 | US-08-882-704A-4 | Sequence 4, Appl |
| 19 | 14.6 | 69.5 | 39 | US-08-411-727-9 | Sequence 9, Appl |
| 20 | 14.6 | 69.5 | 39 | US-08-411-727-17 | Sequence 17, Appl |
| 21 | 14.6 | 69.5 | 728 | US-08-411-727-19 | Sequence 19, Appl |
| 22 | 14.6 | 69.5 | 4097 | US-08-162-809-17 | Sequence 17, Appl |
| 23 | 14.6 | 69.5 | 4097 | US-08-162-809-11 | Sequence 11, Appl |
| 24 | 14.6 | 69.5 | 1146 | US-09-256-496-1 | Sequence 1, Appl |
| 25 | 14.4 | 68.6 | 1457 | US-08-233-389C-2 | Sequence 2, Appl |
| 26 | 14.4 | 68.6 | 1457 | US-08-801-863-2 | Sequence 2, Appl |
| 27 | 14.4 | 68.6 | 1457 | US-08-801-863-2 | Sequence 2, Appl |

| | | | | | | |
|----|------|------|-------|---|------------------|-------------------|
| 28 | 14.4 | 68.6 | 1457 | 2 | US-08-486-596A-2 | Sequence 2, Appl |
| 29 | 14.4 | 68.6 | 1457 | 2 | US-09-004-713-2 | Sequence 2, Appl |
| 30 | 14.4 | 68.6 | 1942 | 4 | US-09-313-300-3 | Sequence 3, Appl |
| 31 | 14.4 | 68.6 | 2173 | 1 | US-08-795-788-1 | Sequence 1, Appl |
| 32 | 14.4 | 68.6 | 2173 | 5 | PCT-US93-03942-1 | Sequence 1, Appl |
| 33 | 14.4 | 68.6 | 3870 | 1 | US-08-138-641-1 | Sequence 1, Appl |
| 34 | 14.4 | 68.6 | 3870 | 1 | US-08-138-133-1 | Sequence 1, Appl |
| 35 | 14.4 | 68.6 | 3893 | 1 | US-08-138-641-3 | Sequence 3, Appl |
| 36 | 14.4 | 68.6 | 3893 | 1 | US-08-138-133-3 | Sequence 3, Appl |
| 37 | 14.4 | 68.6 | 3844 | 4 | US-08-462-437-30 | Sequence 30, Appl |
| 38 | 14.4 | 68.6 | 13104 | 4 | US-08-256-799-4 | Sequence 4, Appl |
| 39 | 14.4 | 68.6 | 13104 | 4 | US-08-462-437-4 | Sequence 4, Appl |
| 40 | 14.2 | 67.6 | 60 | 3 | US-08-584-031-4 | Sequence 4, Appl |
| 41 | 14.2 | 67.6 | 60 | 3 | US-08-780-496-4 | Sequence 4, Appl |
| 42 | 14.2 | 67.6 | 390 | 3 | US-08-584-031-3 | Sequence 3, Appl |
| 43 | 14.2 | 67.6 | 390 | 3 | US-08-780-496-3 | Sequence 3, Appl |
| 44 | 14.2 | 67.6 | 797 | 2 | US-08-969-106-7 | Sequence 7, Appl |
| 45 | 14.2 | 67.6 | 1042 | 3 | US-08-584-031-2 | Sequence 2, Appl |

ALIGNMENTS

RESULT 1
US-08-480-917-12
Sequence 12, Application US/08480917
Patent No. 5820864
GENERAL INFORMATION:
APPLICANT: PARANHOS-BACCALA, Glaucia
APPLICANT: LESENECHAL, Mylene
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: NEW TRYPAOSOMA CRUZI ANTIGEN, AND GENE
TITLE OF INVENTION: ENCODING THE LATTER, THEIR APPLICATION TO THE DETECTION OF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Oliff & Berridge
CITY: Alexandria
STREET: 700 South Washington Street, Suite 300
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,917
FILING DATE: 07-JUN-1995
CLASSIFICATION: A35
ATTORNEY/AGENT INFORMATION:
NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPA 36400
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-480-917-12
Query Match 100.0%; Score 21; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 cgaagagacatgacaaactt 21

Db 1 CGAAGAGACCATGACAACTT 21
|||||
RESULT 2
US-08-480-917-1
; Sequence 1, Application US/08480917
; Patent No. 5820864
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACALA, Glaucia
; APPLICANT: LESENECHAL, Mylene
; APPLICANT: JOUVET, Michel
; TITLE OF INVENTION: NEW TRYANOSOMA CRUZI ANTIGEN, AND GENE
; TITLE OF INVENTION: ENCODING THE LATTER; THEIR APPLICATION TO THE DETECTION OF
; TITLE OF INVENTION: CHAGAS DISEASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,917
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: MPB 36400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3402 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-480-917-1

Query Match 100.0%; Score 21; DB 1; Length 3402;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGAAGAGACCATGACAACTT 21
|||||
Db 1997 CGAAGAGACCATGACAACTT 2017

RESULT 3
US-08-933-750C-69
; Sequence 69, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2029 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAUTY03
; CLONE: 864683
; US-08-933-750C-69

Query Match 80.0%; Score 16.8; DB 2; Length 2029;
Best Local Similarity 90.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GAAGAGACCATGACAACTT 21
|||||
Db 1644 GAAGAGACCATGACAACTT 1663

RESULT 4
US-09-234-613-69
; Sequence 69, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/234.613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/933.750
FILING DATE: September 23, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 2029 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAITFUT03
CLONE: 864683
US-09-234-613-69

Query Match 80.0%; Score 16.8; DB 3; Length 2029;
Best Local Similarity 90.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 gagagaccatgaacaactt 21
|||||
Db 1644 GAGAGACCTGACACACTT 1663

RESULT 5
US-08-861-745B-2
Sequence 2, Application US/08861745B
Patent No. 6165733
GENERAL INFORMATION:
APPLICANT: Cen, Hui
APPLICANT: Williams, Lewis
TITLE OF INVENTION: Gamma II Adaptor
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff
STREET: 1001 G Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/861.745B
FILING DATE: 22-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A
REGISTRATION NUMBER: 32141
REFERENCE/DOCKET NUMBER: 02441.05336
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299

TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2533 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-861-745B-2

Query Match 73.3%; Score 15.4; DB 4; Length 2533;
Best Local Similarity 94.1%; Pred. No. 72;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 gagaccatgaacaactt 21
|||||
Db 823 GAGACCATGACGACTT 839

RESULT 6
US-08-361-467B-4/C
Sequence 4, Application US/08361467B
Patent No. 5633441
GENERAL INFORMATION:
APPLICANT: De Greef, Willy
APPLICANT: Van Emmelo, John
APPLICANT: De Oliveira, Dulce E.
APPLICANT: De Souza, Maria-Helena
TITLE OF INVENTION: PLANTS WITH MODIFIED FLOWERS, SEEDS OR
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/361.467B
FILING DATE: 22-DEC-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/681,492
FILING DATE: 04-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/EP90/01275
FILING DATE: 01-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 89 402 224.3
FILING DATE: 04-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Schultman, Robert M.
REGISTRATION NUMBER: 31,196
REFERENCE/DOCKET NUMBER: 010830-027
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1046 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
IMMEDIATE SOURCE:
CLONE: 3C9

; Sequence 1, Application US/08360673
; Patent No. 5679544
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
; TITLE OF INVENTION: PREPARATION AND USE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,673
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR93/00623
; FILING DATE: 23-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/07785
; FILING DATE: 25-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST92040-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1685 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Kluveromyces lactis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1683
; OTHER INFORMATION: /product= "Protease B gene"
; OTHER INFORMATION: /gene= "K1.PRBI"
US-08-360-673-1

Query Match 72.4%; Score 15.2; DB 1; Length 1685;
Best Local Similarity 85.0%; Pred. No. 86;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 gaagagaccatgaacaactt 21
||||| ||| ||||| |||||
Db 1486 GAAGATACCGTGAAGTACTT 1505

RESULT 10
5206163-2
; Patent No. 5206163
; APPLICANT: RENARD, ANDRE;DINA, DINO;MARTIAL, JOSEPH
; TITLE OF INVENTION: DNA ENCODING BOVINE DIARRHEA
; VIRUS PROTEIN
; NUMBER OF SEQUENCES: 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/550,816

; FILING DATE: 06-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 331,037
; FILING DATE: 29-MAR-1989
; APPLICATION NUMBER: 752,981
; FILING DATE: 08-JUL-1985
; SEQ ID NO: 2;
; LENGTH: 12492
5206163-2

Query Match 72.4%; Score 15.2; DB 6; Length 12492;
Best Local Similarity 85.0%; Pred. No. 1,1e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 gaagagaccatgaacaactt 21
||||| ||| ||||| |||||
Db 12158 gaagagaccatgaacaactt 12177

RESULT 11
US-08-826-267-1
; Sequence 1, Application US/08826267
; Patent No. 5994070
; GENERAL INFORMATION:
; APPLICANT: Streuli, Michel
; TITLE OF INVENTION: No. 5994070e1 TRIO Molecules and Uses Related Thereto
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,267
; FILING DATE: 1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/014,214
; FILING DATE: 27 MARCH (1996)
; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFW-010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-5941
; TELEFAX: (617)227-7400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8906 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 67..8647
US-08-826-267-1

Query Match 71.4%; Score 15; DB 2; Length 8906;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 gaccatgaacaactt 21
||||| ||||| ||||| |||||

Db 2397 GACCATGACAACTT 2411

RESULT 12

PCT-US92-01015-12/C
Sequence 12, Application PC/TUS9201015

GENERAL INFORMATION:

APPLICANT: Davis, Geneva R

APPLICANT: Provow, Sally P

TITLE OF INVENTION: Production of Human Serum Albumin in

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fitch, Even, Tabin & Flannery

STREET: 135 South LaSalle Street, Suite 900

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US92/01015

FILING DATE: 19920204

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/650,040

FILING DATE: 04-FEB-1991

ATTORNEY/AGENT INFORMATION:

NAME: Seidman, Stephanie L.

REGISTRATION NUMBER: 33,779

REFERENCE/DOCKET NUMBER: 50857PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619)552-1311

TELEFAX: (619)552-0095

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 69 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

PCT-US92-01015-12

Query Match 70.5%; Score 14.8; DB 5; Length 69;

Best Local Similarity 88.9%; Pred. No. 86;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 gaagagacatgaacaac 19

Db 26 GAAGAGACAAGAACAC 9

RESULT 13

PCT-US92-01015-3/C

Sequence 3, Application PC/TUS9201015

GENERAL INFORMATION:

APPLICANT: Davis, Geneva R

APPLICANT: Provow, Sally P

TITLE OF INVENTION: Production of Human Serum Albumin in

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fitch, Even, Tabin & Flannery

STREET: 135 South LaSalle Street, Suite 900

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US92/01015

FILING DATE: 19920204

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/650,040

FILING DATE: 04-FEB-1991

ATTORNEY/AGENT INFORMATION:

NAME: Seidman, Stephanie L.

REGISTRATION NUMBER: 33,779

REFERENCE/DOCKET NUMBER: 50857PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619)552-1311

TELEFAX: (619)552-0095

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 72 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..72

OTHER INFORMATION:

PCT-US92-01015-3

Query Match 70.5%; Score 14.8; DB 5; Length 72;

Best Local Similarity 88.9%; Pred. No. 87;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 gaagagacatgaacaac 19

Db 44 GAAGAGACAAGAACAC 27

RESULT 14

US-08-026-145-3/C

Sequence 3, Application US/08026145

Patent No. 5378614

GENERAL INFORMATION:

APPLICANT: Petersen, Jens G. Iltse

TITLE OF INVENTION: Method for Making TPPI Analogues

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 5378614 of No. 5378614th America, Inc.

STREET: 405 Lexington Avenue, Suite 6200

CITY: New York

STATE: N. Y.

COUNTRY: United States of America

ZIP: 10174-6201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/026,145

FILING DATE: 19930302

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/828,920

FILING DATE: 27-JAN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/DK/90/00212

FILING DATE: 17-AUG-1990

PRIOR APPLICATION DATA:
APPLICATION NUMBER: DK 4080/89
FILING DATE: 18-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Agtis, Cheryl H.
REGISTRATION NUMBER: 34086
REFERENCE/DOCKET NUMBER: 3321.214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 81 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Synthetic
FEATURE:
NAME/KEY: CDS
LOCATION: 1..81
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..54
FEATURE:
NAME/KEY: misc_feature
LOCATION: 55..72
US-08-026-145-3

Query Match 70.5%; Score 14.8; DB 1; Length 81;
Best Local Similarity 88.9%; Pred. No. 88;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 gaagagaccatgaacaac 19
||||| || |||||
DB 44 GAAGAGAACAGAACAC 27

RESULT 15
US-07-872-673B-4/C
Sequence 4, Application US/07872673B
Patent No. 5578466
GENERAL INFORMATION:
APPLICANT: Toshiya HAYANO, Setsuko KATO, No. 5578466uhiro TAKAHASHI, and Masanori
TITLE OF INVENTION: Co-expression system of Protein Disulfide Isomerase Gene an
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS & CUSHMAN
STREET: 130 WATER STREET
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Type 2DD, 3.50 inch, 720 KB
COMPUTER: Apple Macintosh SE
OPERATING SYSTEM: Apple DOS
SOFTWARE: Microsoft Word Version 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/872,673B
FILING DATE: 19920417
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: Japanese Patent Application No. 5578466, 114074/91 and 31160
FILING DATE: 18-APR-1991 and 30-OCT-1991
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 96
TYPE: NUCLEIC ACID

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid
US-07-872-673B-4

Query Match 70.5%; Score 14.8; DB 1; Length 96;
Best Local Similarity 88.9%; Pred. No. 90;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 gaagagaccatgaacaac 19
||||| || |||||
DB 44 GAAGAGAACAGAACAC 27

Search completed: September 21, 2001, 02:16:45
Job time: 28238 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 02:20:06 : Search time 6407.28 Seconds
(without alignments)
48.431 Million cell updates/sec

Title: US-09-138-735-12

Perfect score: 21
Sequence: 1 cgaagagacatgaacactt 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 14155048 seqs, 7388405095 residues

Total number of hits satisfying chosen parameters: 28310096

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1: /cgnl_7/ptodata/1/pna/PCRNUS.COMB.seq:*
2: /cgnl_7/ptodata/1/pna/US06.COMB.seq:*
3: /cgnl_7/ptodata/1/pna/US07.COMB.seq:*
4: /cgnl_7/ptodata/1/pna/US08.COMB.seq:*
5: /cgnl_7/ptodata/1/pna/US081.COMB.seq:*
6: /cgnl_7/ptodata/1/pna/US082.COMB.seq:*
7: /cgnl_7/ptodata/1/pna/US083.COMB.seq:*
8: /cgnl_7/ptodata/1/pna/US084.COMB.seq:*
9: /cgnl_7/ptodata/1/pna/US085.COMB.seq:*
10: /cgnl_7/ptodata/1/pna/US086.COMB.seq:*
11: /cgnl_7/ptodata/1/pna/US087.COMB.seq:*
12: /cgnl_7/ptodata/1/pna/US088.COMB.seq:*
13: /cgnl_7/ptodata/1/pna/US089.COMB.seq:*
14: /cgnl_7/ptodata/1/pna/US090.COMB.seq:*
15: /cgnl_7/ptodata/1/pna/US091.COMB.seq:*
16: /cgnl_7/ptodata/1/pna/US092.COMB.seq:*
17: /cgnl_7/ptodata/1/pna/US093.COMB.seq:*
18: /cgnl_7/ptodata/1/pna/US094.COMB.seq:*
19: /cgnl_7/ptodata/1/pna/US095.COMB.seq:*
20: /cgnl_7/ptodata/1/pna/US095b.COMB.seq:*
21: /cgnl_7/ptodata/1/pna/US095c.COMB.seq:*
22: /cgnl_7/ptodata/1/pna/US095d.COMB.seq:*
23: /cgnl_7/ptodata/1/pna/US096a.COMB.seq:*
24: /cgnl_7/ptodata/1/pna/US096b.COMB.seq:*
25: /cgnl_7/ptodata/1/pna/US096c.COMB.seq:*
26: /cgnl_7/ptodata/1/pna/US096d.COMB.seq:*
27: /cgnl_7/ptodata/1/pna/US096e.COMB.seq:*
28: /cgnl_7/ptodata/1/pna/US097a.COMB.seq:*
29: /cgnl_7/ptodata/1/pna/US097b.COMB.seq:*
30: /cgnl_7/ptodata/1/pna/US097c.COMB.seq:*
31: /cgnl_7/ptodata/1/pna/US098.COMB.seq:*
32: /cgnl_7/ptodata/1/pna/US6000.COMB.seq:*
33: /cgnl_7/ptodata/1/pna/US6001.COMB.seq:*
34: /cgnl_7/ptodata/1/pna/US6002.COMB.seq:*
35: /cgnl_7/ptodata/1/pna/US6003.COMB.seq:*
36: /cgnl_7/ptodata/1/pna/US6004.COMB.seq:*
37: /cgnl_7/ptodata/1/pna/US6005.COMB.seq:*
38: /cgnl_7/ptodata/1/pna/US6006.COMB.seq:*
39: /cgnl_7/ptodata/1/pna/US6007.COMB.seq:*
40: /cgnl_7/ptodata/1/pna/US6008.COMB.seq:*
41: /cgnl_7/ptodata/1/pna/US6009.COMB.seq:*
42: /cgnl_7/ptodata/1/pna/US6010.COMB.seq:*
43: /cgnl_7/ptodata/1/pna/US6011.COMB.seq:*

44: /cgnl_7/ptodata/1/pna/US6012.COMB.seq:*
45: /cgnl_7/ptodata/1/pna/US6013.COMB.seq:*
46: /cgnl_7/ptodata/1/pna/US6014.COMB.seq:*
47: /cgnl_7/ptodata/1/pna/US6015.COMB.seq:*
48: /cgnl_7/ptodata/1/pna/US6016.COMB.seq:*
49: /cgnl_7/ptodata/1/pna/US6017.COMB.seq:*
50: /cgnl_7/ptodata/1/pna/US6018.COMB.seq:*
51: /cgnl_7/ptodata/1/pna/US6019.COMB.seq:*
52: /cgnl_7/ptodata/1/pna/US6020.COMB.seq:*
53: /cgnl_7/ptodata/1/pna/US6021.COMB.seq:*
54: /cgnl_7/ptodata/1/pna/US6022.COMB.seq:*
55: /cgnl_7/ptodata/1/pna/US6023.COMB.seq:*
56: /cgnl_7/ptodata/1/pna/US6024.COMB.seq:*
57: /cgnl_7/ptodata/1/pna/US6025.COMB.seq:*
58: /cgnl_7/ptodata/1/pna/US6026.COMB.seq:*
59: /cgnl_7/ptodata/1/pna/US6027.COMB.seq:*
60: /cgnl_7/ptodata/1/pna/US6028.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-----------------------|
| 1 | 21 | 100.0 | 21 | 13 | US-08-988-242-7 |
| 2 | 21 | 100.0 | 21 | 15 | US-09-138-736-12 |
| 3 | 21 | 100.0 | 3402 | 13 | US-08-988-242-1 |
| 4 | 21 | 100.0 | 3402 | 15 | US-09-138-736-1 |
| 5 | 18 | 85.7 | 13837 | 56 | US-60-248-505-93 |
| 6 | 18 | 85.7 | 332768 | 54 | US-60-229-515-89 |
| 7 | 18 | 85.7 | 332768 | 54 | US-60-229-515-90 |
| 8 | 18 | 85.7 | 32768 | 54 | US-60-229-515-91 |
| 9 | 18 | 85.7 | 32768 | 54 | US-60-229-515-92 |
| 10 | 18 | 85.7 | 34117 | 56 | US-60-245-225-193 |
| 11 | 17.8 | 84.8 | 240 | 21 | US-08-917-047-472 |
| 12 | 17.8 | 84.8 | 240 | 21 | US-09-540-212a-32930 |
| 13 | 17.8 | 84.8 | 240 | 34 | US-60-025-203-472 |
| 14 | 17.8 | 84.8 | 2206 | 58 | US-60-268-380-121 |
| 15 | 17.4 | 82.9 | 278 | 19 | US-09-521-640-66611 |
| 16 | 17.4 | 82.9 | 286 | 21 | US-09-540-229-150778 |
| 17 | 17.4 | 82.9 | 477 | 15 | US-09-107-532-323 |
| 18 | 17.4 | 82.9 | 524 | 40 | US-60-082-300-21418 |
| 19 | 17.4 | 82.9 | 687 | 18 | US-09-417-507-21543 |
| 20 | 17.4 | 82.9 | 969 | 15 | US-09-107-532-258 |
| 21 | 17.4 | 82.9 | 1044 | 16 | US-08-270-849B-149102 |
| 22 | 16.8 | 80.0 | 195 | 11 | US-08-734-050-589 |
| 23 | 16.8 | 80.0 | 195 | 13 | US-08-992-868-589 |
| 24 | 16.8 | 80.0 | 195 | 21 | US-09-540-764-27960 |
| 25 | 16.8 | 80.0 | 195 | 32 | US-60-006-111-589 |
| 26 | 16.8 | 80.0 | 196 | 12 | US-08-808-904-4845 |
| 27 | 16.8 | 80.0 | 196 | 12 | US-08-808-904A-4845 |
| 28 | 16.8 | 80.0 | 196 | 21 | US-09-540-212a-26524 |
| 29 | 16.8 | 80.0 | 233 | 14 | US-09-013-812-2028 |
| 30 | 16.8 | 80.0 | 233 | 21 | US-09-540-764-34474 |
| 31 | 16.8 | 80.0 | 233 | 35 | US-60-036-570-2028 |
| 32 | 16.8 | 80.0 | 236 | 11 | US-08-733-814-81 |
| 33 | 16.8 | 80.0 | 236 | 4 | US-09-540-764-16183 |
| 34 | 16.8 | 80.0 | 237 | 12 | US-08-857-212-913 |
| 35 | 16.8 | 80.0 | 237 | 18 | US-09-480-902-5904 |
| 36 | 16.8 | 80.0 | 237 | 12 | US-09-480-902-5904 |
| 37 | 16.8 | 80.0 | 237 | 21 | US-09-540-208-29228 |
| 38 | 16.8 | 80.0 | 237 | 21 | US-09-540-229-88357 |
| 39 | 16.8 | 80.0 | 239 | 15 | US-09-107-910-5136 |
| 40 | 16.8 | 80.0 | 239 | 21 | US-09-540-229-68062 |
| 41 | 16.8 | 80.0 | 241 | 14 | US-09-074-999-1854 |
| 42 | 16.8 | 80.0 | 241 | 36 | US-60-048-431-1854 |
| 43 | 16.8 | 80.0 | 243 | 13 | US-08-903-474-1860 |
| 44 | 16.8 | 80.0 | 243 | 21 | US-09-540-208-1064 |
| 45 | 16.8 | 80.0 | 253 | 13 | US-08-901-904-3120 |

ALIGNMENTS

```
RESULT 1
US-08-988-242-7
; Sequence 7, Application US/08988242
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: LESENECHAL, MYLENE
; APPLICANT: JOLIVET, MICHEL
; APPLICANT: MANDRAND, BERNARD
; TITLE OF INVENTION: NEW TRIPANOSOMA CRUZI ANTIGEN, GENE
; TITLE OF INVENTION: ENCODING THEREFOR, AND METHODS OF DETECTING AND TREATING
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/988,242
; FILING DATE: 10-DEC-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: MPB 36400A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-988-242-7

Query Match          100.0%; Score 21; DB 13; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cgaagagaccatgaacaactt 21
DB 1 CGAAGAGACCATGACAACACTT 21

RESULT 2
US-09-138-736-12
; Sequence 12, Application US/09138736
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: LESENECHAL, MYLENE
; APPLICANT: JOLIVET, MICHEL
; TITLE OF INVENTION: NEW TRIPANOSOMA CRUZI ANTIGEN, AND GENE
; TITLE OF INVENTION: ENCODING THE LATTER, THEIR APPLICATION TO THE DETECTION OF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
```

```
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/138,736
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,917
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: MPB 36400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-138-736-12

Query Match          100.0%; Score 21; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cgaagagaccatgaacaactt 21
DB 1 CGAAGAGACCATGACAACACTT 21

RESULT 3
US-08-988-242-1
; Sequence 1, Application US/08988242
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: LESENECHAL, MYLENE
; APPLICANT: JOLIVET, MICHEL
; APPLICANT: MANDRAND, BERNARD
; TITLE OF INVENTION: NEW TRIPANOSOMA CRUZI ANTIGEN, GENE
; TITLE OF INVENTION: ENCODING THEREFOR, AND METHODS OF DETECTING AND TREATING
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/988,242
; FILING DATE: 10-DEC-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: MPB 36400A
```

TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3402 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-988-242-1

Query Match 100.0%; Score 21; DB 13; Length 3402;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cgaagagacatgaacaact 21
|||||
Db 1997 CGAAGACCATGAACTT 2017

RESULT 4
US-09-138-736-1

Sequence 1, Application US/09138736
GENERAL INFORMATION:
APPLICANT: PARANHOS-BACCALA, Glaucia
APPLICANT: LERECHAL, Mylene
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: NEW TRYPANOSOMA CRUZI ANTIGEN, AND GENE
TITLE OF INVENTION: ENCODING THE LATTER, THEIR APPLICATION TO THE DETECTION OF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSER: Oliff & Berridge
STREET: 700 South Washington Street, Suite 300
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/138,736
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/480,917
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: MPB 36400
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3402 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-138-736-1

Query Match 100.0%; Score 21; DB 15; Length 3402;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cgaagagacatgaacaact 21
|||||
Db 1997 CGAAGACCATGAACTT 2017

RESULT 5
US-60-248-505-93

Sequence 93, Application US/60248505
GENERAL INFORMATION:
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
FILE REFERENCE: C1000918
CURRENT APPLICATION NUMBER: US/60/248,505
CURRENT FILING DATE: 2000-11-15
NUMBER OF SEQ ID NOS: 1998
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 93
LENGTH: 13837
TYPE: DNA
ORGANISM: human
US-60-248-505-93

Query Match 85.7%; Score 18; DB 56; Length 13837;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 aagagacatgaacaact 20
|||||
Db 772 aagagacatgaacaact 789

RESULT 6
US-60-229-515-89/c

Sequence 89, Application US/60229515
GENERAL INFORMATION:
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
FILE REFERENCE: C1000776
CURRENT APPLICATION NUMBER: US/60/229,515
CURRENT FILING DATE: 2000-09-05
NUMBER OF SEQ ID NOS: 2013
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 89
LENGTH: 32768
TYPE: DNA
ORGANISM: HUMAN
US-60-229-515-89

Query Match 85.7%; Score 18; DB 54; Length 32768;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 aagagacatgaacaact 20
|||||
Db 4732 AAGAGACCATGAACTT 4715

RESULT 7
US-60-229-515-90/c

Sequence 90, Application US/60229515
GENERAL INFORMATION:
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
FILE REFERENCE: C1000776
CURRENT APPLICATION NUMBER: US/60/229,515

```

; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 2013
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 90
; LENGTH: 32768
; TYPE: DNA
; ORGANISM: HUMAN
US-60-229-515-90

```

```

Query Match      85.7%: Score 18; DB 54; Length 32768;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 3 aagagaccatgaacaact 20
|||||
Db 4732 AAGAGACCATGAACAAC 4715

```

```

RESULT 8
US-60-229-515-91/C
; Sequence 91, Application US/60229515
; GENERAL INFORMATION:

```

```

; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: CL000776
; CURRENT APPLICATION NUMBER: US/60/229,515
; NUMBER OF SEQ ID NOS: 2000-09-05
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91
; LENGTH: 32768
; TYPE: DNA
; ORGANISM: HUMAN
US-60-229-515-91

```

```

US-60-229-515-91

```

```

Query Match      85.7%: Score 18; DB 54; Length 32768;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 3 aagagaccatgaacaact 20
|||||
Db 4732 AAGAGACCATGAACAAC 4715

```

```

RESULT 9
US-60-229-515-92/C
; Sequence 92, Application US/60229515
; GENERAL INFORMATION:

```

```

; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: CL000776
; CURRENT APPLICATION NUMBER: US/60/229,515
; NUMBER OF SEQ ID NOS: 2000-09-05
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 92
; LENGTH: 32768
; TYPE: DNA
; ORGANISM: HUMAN
US-60-229-515-92

```

```

Query Match      85.7%: Score 18; DB 54; Length 32768;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 3 aagagaccatgaacaact 20

```

```

Db 4732 AAGAGACCATGAACAAC 4715
|||||

```

```

RESULT 10
US-60-245-225-193
; Sequence 193, Application US/60245225
; GENERAL INFORMATION:

```

```

; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: CL000885
; CURRENT APPLICATION NUMBER: US/60/245,225
; NUMBER OF SEQ ID NOS: 2000-11-03
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 193
; LENGTH: 34117
; TYPE: DNA
; ORGANISM: Human
US-60-245-225-193

```

```

Query Match      85.7%: Score 18; DB 56; Length 34117;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 3 aagagaccatgaacaact 20
|||||
Db 773 aagagaccatgaacaact 790

```

```

RESULT 11
US-08-917-047-472/C
; Sequence 472, Application US/08917047
; GENERAL INFORMATION:

```

```

; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Deleage, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN FETAL LUNG
; NUMBER OF SEQUENCES: 2077
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/917,047
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C., Ph.D.
; REGISTRATION NUMBER: 39132
; REFERENCE/DOCKET NUMBER: PD-0230P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 472:
; SEQUENCE CHARACTERISTICS:

```

```

;
; LENGTH: 240 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: cdna
;
; IMMEDIATE SOURCE:
;
; CLONE: 1849638
;
;
US-08-917-047-472

```

| | | | | |
|-----------------------|-----------------|---------------|-----------|-------------|
| Query Match | 84.8% | Score 17.8; | DB 13. | Length 240; |
| Best Local Similarity | 90.5% | Pred. No. 2 | 6e+02; | |
| Matches 19; | Conservative 0; | Mismatches 2; | Indels 0; | Gaps 0 |

| | | | |
|----|-----|-----------------------|----|
| QY | 1 | cgaagagaccatgaacact | 21 |
| | | | |
| Db | 115 | CTAAGAGACCATGAGCAACTT | 95 |

```

RESULT 12
US-09-540-212A-32930/C
: Sequence 32930, Application US/09540212A
: GENERAL INFORMATION:
: APPLICANT: Seilhamer, Jeffrey J.
: APPLICANT: Delegeane, Angelo M.
: APPLICANT: Stuart, Susan G.
: APPLICANT: Stuve, Laura L.
: APPLICANT: Mullaby, Sara J.
: APPLICANT: Naughton, Rebecca E.
: TITLE OF INVENTION: POLYNUCLEOTIDES OF AIRWAY AND LUNG SYSTEM TISSUE
: FILE REFERENCE: PD-1034 CIP
: CURRENT APPLICATION NUMBER: US/09/540.212A
: CURRENT FILING DATE: 2000-03-31
: NUMBER OF SEQ ID NOS: 67551
: SOFTWARE: PERL Program
: SEQ ID NO 32930
: LENGTH: 240
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc-feature
: OTHER INFORMATION: Incyte ID No: hu00018466
US-09-540-212A-32930

```

| | | | | |
|-----------------------|----------------|------------------|----------|------------|
| Query Match | 84.8% | Score 17.8 | DB 21 | Length 240 |
| Best Local Similarity | 90.5% | Pred No. 2.6e+02 | | |
| Matches 19 | Conservative 0 | Mismatches 2 | Indels 0 | Gaps 0 |

```

QY      1  cgaagagaccatgaacaact 21
          | | | | | | | | | |
Db      115 CTAAGAGACCATGAGCAACTT 95

```

RESULT 13
US-60-025-203-472/C
; Sequence 472, Application US/60025203
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akherblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Kingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN FETAL LUNG
; NUMBER OF SEQUENCES: 2077
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
;

STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/025,203
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C., Ph.D.
REGISTRATION NUMBER: 39132
REFERENCE/DOCKET NUMBER: PD-0230P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 472:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
IMMEDIATE SOURCE:
CLONE: 1849638
US-60-025-203-472

| | | | | |
|--------------------------|-------|--------------------|-----------|-------------|
| Query Match | 84.8% | Score 17.8; | DB 34. | Length 240; |
| Best Local Similarity | 90.5% | Pred. No. 2.6e+02; | | |
| Matches 19; Conservative | 0; | Mismatches 2; | Indels 0; | Gaps 0; |

```

QY      1 cgaagagaccatgaacaactt 21
        | | | | | | | | | | | | |
Db      115 cTAAGAGACCATGAGCAACTT 95

```

```

RESULT 14
US-60-268-380-121
; Sequence 121, Application US/60268380
; GENERAL INFORMATION:
; APPLICANT: Huel-Wei Chen
; TITLE OF INVENTION: GENES EXPRESSED IN SENESENCESCENCE
; FILE REFERENCE: PA-0043 P
; CURRENT APPLICATION NUMBER: US/60/268,380
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PERL Program
; SEQ ID NO 121
; LENGTH: 2206
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 474629.A
; NAME/KEY: unsure
; LOCATION: 486
; OTHER INFORMATION: a, t, c, g, or other
US-60-268-380-121

```

| | | | | |
|-----------------------|-----------------|--------------------|-----------|--------------|
| Query Match | 84.8% | Score 17.8: | DB 58: | Length 2206: |
| Best Local Similarity | 90.5% | Pred. No. 3.7e+02: | | |
| Matches 19: | Conservative 0: | Mismatches 2: | Indels 0: | Gaps 0: |

QY 1 cgaagagaccatgaacaact 21
||||| | ||||| |||||
Db 345 cgaagaatcatgaacaact 365

```

RESULT 15
US-09-521-640-66611
; Sequence 66611, Application US/09521640
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Halling, Conrad H.
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15750)D
; CURRENT APPLICATION NUMBER: US/09/521,640
; CURRENT FILING DATE: 2000-03-10
; NUMBER OF SEQ. ID NOS: 304701
; SEQ ID NO 66611
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: unsure at all n locations
US-09-521-640-66611
    
```

```

Query Match      82.9%; Score 17.4; DB 19; Length 278;
Best Local Similarity 94.7%; Pred. No. 4.1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 gaagagaccatgaacaact 20
   ||||||| |||||||
Db 245 gaagagaccgtgaacaact 263
    
```

Search completed: September 21, 2001, 02:20:15
 Job time: 30487 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 02:40:43 : Search time 1035.34 Seconds
(without alignments)
40.480 Million cell updates/sec

Title: US-09-138-735-12

Perfect score: 21

Sequence: 1 cgaagagaccatgaacaactt 21

Scoring table: IDENTITY_NUC

Searched: 1741792 seqs, 997871973 residues

Total number of hits satisfying chosen parameters: 3483584

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_MA_New : *

- 1: /cgn1_7/ptodata/2/pna/US06_NEW_COMB.seq:*
- 2: /cgn1_7/ptodata/2/pna/US06_NEW_COMB.seq:*
- 3: /cgn1_7/ptodata/2/pna/US07_NEW_COMB.seq:*
- 4: /cgn1_7/ptodata/2/pna/US08_NEW_COMB.seq:*
- 5: /cgn1_7/ptodata/2/pna/US09_NEW_COMB.seq:*
- 6: /cgn1_7/ptodata/2/pna/US09_NEW_COMB.seq:*
- 7: /cgn1_7/ptodata/2/pna/US09_NEW_COMB.seq:*
- 8: /cgn1_7/ptodata/2/pna/US09_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------------|-------------------|
| 1 | 21 | 100.0 | 21 | US-09-138-735-12 | Sequence 12, Appl |
| 2 | 21 | 100.0 | 3402 | US-09-138-735-1 | Sequence 1, Appl |
| 3 | 16.8 | 80.0 | 422 | US-09-866-555-6605 | Sequence 6005, Ap |
| 4 | 16.8 | 80.0 | 470 | US-09-904-809-7574 | Sequence 7574, Ap |
| 5 | 16.8 | 80.0 | 632 | US-09-691-918-480 | Sequence 490, App |
| 6 | 16.8 | 80.0 | 794 | US-09-861-478-6875 | Sequence 6875, Ap |
| 7 | 16.8 | 80.0 | 1107 | US-09-757-034-62 | Sequence 62, Appl |
| 8 | 16.8 | 80.0 | 1187 | US-09-898-888-12899 | Sequence 12899, A |
| 9 | 16.8 | 80.0 | 1931 | US-09-925-297-150 | Sequence 150, App |
| 10 | 16.8 | 80.0 | 2225 | US-09-764-875-151 | Sequence 151, App |
| 11 | 16.8 | 80.0 | 13630 | US-09-764-868-1369 | Sequence 1369, Ap |
| 12 | 16.8 | 80.0 | 13630 | US-09-764-875-1220 | Sequence 1220, Ap |
| 13 | 16.8 | 80.0 | 94378 | US-09-803-736-500 | Sequence 500, App |
| 14 | 16.8 | 80.0 | 105863 | US-09-803-736-1145 | Sequence 1145, Ap |
| 15 | 16.8 | 80.0 | 308503 | US-09-803-110-214 | Sequence 214, App |
| 16 | 16.4 | 78.1 | 369 | US-09-904-809-12595 | Sequence 12595, A |
| 17 | 16.4 | 78.1 | 405 | US-09-773-486-3347 | Sequence 3347, Ap |
| 18 | 16.4 | 78.1 | 852 | US-09-764-864-285 | Sequence 285, App |
| 19 | 16.4 | 78.1 | 852 | US-09-764-891-1407 | Sequence 1407, Ap |
| 20 | 16.4 | 78.1 | 2883 | US-09-14827-2476 | Sequence 2476, Ap |
| 21 | 16.4 | 78.1 | 2883 | US-09-08656-2635 | Sequence 2635, A |
| 22 | 16.4 | 78.1 | 2883 | US-09-08631-17735 | Sequence 17735, A |
| 23 | 16.4 | 78.1 | 2883 | US-09-08631-25966 | Sequence 25966, A |
| 24 | 16.4 | 78.1 | 3659 | US-09-764-864-1668 | Sequence 1668, Ap |
| 25 | 16.4 | 78.1 | 3659 | US-09-764-891-7820 | Sequence 7820, Ap |

| | | | | | |
|----|------|------|--------|---------------------|--------------------|
| 26 | 16.4 | 78.1 | 200576 | US-09-803-736-580 | Sequence 580, App |
| 27 | 16.4 | 78.1 | 206606 | US-09-803-736-581 | Sequence 581, App |
| 28 | 16.2 | 77.1 | 255 | US-60-253-457-6542 | Sequence 6542, App |
| 29 | 16.2 | 77.1 | 308 | US-60-253-833-10885 | Sequence 10885, A |
| 30 | 16.2 | 77.1 | 316 | US-60-253-457-3359 | Sequence 3359, App |
| 31 | 16.2 | 77.1 | 354 | US-60-253-653-8211 | Sequence 8211, App |
| 32 | 16.2 | 77.1 | 430 | US-09-906-555-13496 | Sequence 13496, A |
| 33 | 16.2 | 77.1 | 451 | US-09-864-761-5 | Sequence 5, Appl |
| 34 | 16.2 | 77.1 | 496 | US-60-253-653-12425 | Sequence 12425, A |
| 35 | 16.2 | 77.1 | 536 | US-60-253-653-5302 | Sequence 5302, App |
| 36 | 16.2 | 77.1 | 590 | US-60-253-653-6041 | Sequence 6041, App |
| 37 | 16.2 | 77.1 | 604 | US-60-253-653-4496 | Sequence 4496, App |
| 38 | 16.2 | 77.1 | 605 | US-60-253-653-4652 | Sequence 4652, App |
| 39 | 16.2 | 77.1 | 636 | US-60-253-833-10704 | Sequence 10704, A |
| 40 | 16.2 | 77.1 | 668 | US-60-253-653-3848 | Sequence 3848, App |
| 41 | 16.2 | 77.1 | 999 | US-09-08631-626 | Sequence 626, App |
| 42 | 16.2 | 77.1 | 1218 | US-09-803-110-8105 | Sequence 8105, App |
| 43 | 16.2 | 77.1 | 67757 | US-09-803-736-1252 | Sequence 1252, App |
| 44 | 16.2 | 77.1 | 69506 | US-09-803-736-1258 | Sequence 1258, App |
| 45 | 16.2 | 77.1 | 110826 | US-09-803-736-547 | Sequence 547, App |

ALIGNMENTS

```

RESULT 1
US-09-138-735-12
; Sequence 12, Application US/09138735
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: LESENECHAL, Mylene
; APPLICANT: JOLIVET, Michel
; TITLE OF INVENTION: TRYPANOSOMA CRUZI ANTIGEN, GENE ENCODING THEREFOR AND METHODS
; FILE REFERENCE: WPB 36400B
; CURRENT FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: US 08/480,917
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: FR 94/10132
; PRIOR FILING DATE: 1994-08-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: probe/primer
US-09-138-735-12

Query Match 100.0%; Score 21; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cgaagagaccatgaacaactt 21
Db 1 cgaagagaccatgaacaactt 21

RESULT 2
US-09-138-735-1
; Sequence 1, Application US/09138735
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: LESENECHAL, Mylene
; APPLICANT: JOLIVET, Michel
; TITLE OF INVENTION: TRYPANOSOMA CRUZI ANTIGEN, GENE ENCODING THEREFOR AND METHODS
; FILE REFERENCE: WPB 36400B
; CURRENT APPLICATION NUMBER: US/09138735
; CURRENT FILING DATE: 1998-08-24

```

? PRIOR APPLICATION NUMBER: US 08/480,911
 ? PRIOR FILING DATE: 1995-06-07
 ? PRIOR APPLICATION NUMBER: FR 94/10132
 ? PRIOR FILING DATE: 1994-08-12
 ? NUMBER OF SEQ ID NOS: 13
 ? SOFTWARE: PatentIn version 3.0
 ? SEQ ID NO: 1
 ? LENGTH: 3402
 ? TYPE: DNA
 ? ORGANISM: Trypanosoma cruzi
 ? US-09-138-735-1

| | | | | |
|--------------------------|---------|----------------|-------|-------------------|
| Query Match | 100.0%; | Score 21; | DB 6; | Length 3402; |
| Best Local Similarity | 100.0%; | Pred. No. 1.2; | | |
| Matches 21; Conservative | 0; | Mismatches | 0; | Indels 0; Gaps 0; |

| Qy | 1 | cgaaagagaccatgaacaact | 21 |
|----|------|-----------------------|------|
| Db | 1997 | cgaaagagaccatgaacaact | 2017 |

```

RESULT      3
US-09-866-555-6005/C
; Sequence 6005, Application US/09866555
; GENERAL INFORMATION:
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaltanya S.
; APPLICANT: Algate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Klee, Jennifer
; TITLE OF INVENTION: LUNG TUMOR-ASSOCIATED SEQUENCES AND METHODS OF USE
; FILL REFERENCE: 210121.580
; CURRENT APPLICATION NUMBER: US/09/866.555
; NUMBER OF SEQ ID NOS: 20487
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6005
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-866-555-6005

```

| | | | | |
|--------------------------|-------|---------------|-----------|-------------|
| Query Match | 80.0% | Score 16.8; | DB 6; | Length 422; |
| Best Local Similarity | 90.0% | Pred. No. 98; | | |
| Matches 18; Conservative | 0; | Mismatches 2; | Indels 0; | Gaps 0 |

| | | | |
|----|-----|---------------------|-----|
| Qy | 2 | gaagagaccatgacaactt | 21 |
| | | | |
| Db | 316 | GAAGGACCATGAACAACGT | 297 |

```

RESULT      4
US-09-904-809-7574/C
: Sequence 7574, Application US/09904809
:
GENERAL INFORMATION:
:
APPLICANT: Hysqg, Inc.
:
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
:
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
:
FILE REFERENCE: 20411-757
:
CURRENT APPLICATION NUMBER: US/09/904, 809
:
CURRENT FILING DATE: 2001-07-12
:
PRIORITY FILING DATE: 1999-01-22
:
PRIORITY FILING DATE: 1999-01-22
:
NUMBER OF SEQ ID NOS: 21025
:
SOFTWARE: FastSeq.for Windows Version 3.0
:
SEQ ID NO 7574
:
LENGTH: 470
:
TYPE: DNA
:
ORGANISM: Homo sapiens
:
FEATURE:

```

```

; NAME/KEY: misc_feature
; LOCATION: (1)...(470)
; OTHER INFORMATION: n = A,T,C or G
US-09-904-809-7574

```

| | | | | |
|--------------------------|--------|------------------|-----------|-------------|
| Query Match | 80.0%; | Score 16.8; | DB 6; | Length 470; |
| Best Local Similarity | 90.0%; | Pred. No. 1e+02; | | |
| Matches 18; Conservative | 0; | Mismatches 2; | Indels 0; | Gaps 0; |

```

Qy      2 gaagagaccatgaacaact 21
          ||| ||| ||| ||| ||| ||| |||
Db     329 GAAGAGACCTTGAACAAGTT 310

```

```

RESULT      5
US-09-691-918-490/c
: Sequence 490, Application US/09691918
: GENERAL INFORMATION:
: APPLICANT: Gueglert, Karl J.
: APPLICANT: Doyle, Martin
: APPLICANT: Leader, David
: APPLICANT: Momiyama, Monika G.
: APPLICANT: Mooney, Mark
: APPLICANT: Schuch, Wolfgang
: APPLICANT: Didler, Thomas
: APPLICANT: Thomas, Paul
: TITLE OF INVENTION: WHAT METABOLIC PROTEINS
: FILE REFERENCE: PL-0034 P
: CURRENT APPLICATION NUMBER: US/09/691,918
: CURRENT FILING DATE: 1999-09-16
: NUMBER OF SEQ ID NOS: 910
: SOFTWARE: PERL Program
: SEQ ID NO 490
: LENGTH: 632
: TYPE: DNA
: ORGANISM: Triticum aestivum
: FEATURE:
: NAME/KEY: unsure
: LOCATION: 23, 53, 72
: OTHER INFORMATION: a or g or c or t, unknown, or other
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte template ID No: 014558.1c
US-09-691-918-490

```

| | | | | |
|-----------------------|-----------------|------------------|-----------|-------------|
| Query Match | 80.0%; | Score 16.8; | DB 6; | length 632; |
| Best Local Similarity | 90.0%; | Pred. No. 1e+02; | | |
| Matches 18; | Conservative 0; | Mismatches 2; | Indels 0; | Gaps 0 |

Qy 2 gaagagaccatgacaactt 21
||||| ||| ||||| |||
Db 539 GAAGAAACGATGAACAACTT 520

```

RESULT      6
US-09-661-478-6875/c
: Sequence 6875, Application US/09861478
: GENERAL INFORMATION:
: APPLICANT: Bush, David F.
: APPLICANT: Chasalow, Scott D.
: APPLICANT: Epp, J. Andrew
: APPLICANT: Hauge, Brian M.
: APPLICANT: Lai, Chao-Qiang
: APPLICANT: Laurie, Cathy C.
: APPLICANT: McCarroll, Robert M.
: TITLE OF INVENTION: Maize Polymorphisms and Methods of Genotyping
: FILE REFERENCE: 38-1052018/A
: CURRENT APPLICATION NUMBER: US/09/861,478
: CURRENT FILING DATE: 2001-05-18
: NUMBER OF SEQ ID NOS: 10380
: SEQ ID NO 6875
: LENGTH: 794

```


Db 1674 gaagagacctgacaagctt 1693

RESULT 11

```
US-09-764-868-1369
; Sequence 1369, Application US/09764868
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P7232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1369
; LENGTH: 13630
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (2408)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2615)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2616)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2635)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2636)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2639)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2642)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2643)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2649)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2650)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2652)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2653)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (9943)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-1369
```

Query Match 80.0%; Score 16.8; DB 6; Length 13630;
Best Local Similarity 90.0%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 gaagagacctgacaagctt 21
Db 13109 gaagagacctgacaagctt 13128

RESULT 12

US-09-764-875-1220

```
; Sequence 1220, Application US/09764875
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P7202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1220
; LENGTH: 13630
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (2408)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2615)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2616)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2635)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2636)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2639)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2642)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2643)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2649)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2650)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2652)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2653)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (9943)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-875-1220
```

Query Match 80.0%; Score 16.8; DB 6; Length 13630;
Best Local Similarity 90.0%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 gaagagacctgacaagctt 21
Db 13109 gaagagacctgacaagctt 13128

RESULT 13
US-09-803-736-500/c
; Sequence 500, Application US/09803736
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Levin, Irena M.

```
; APPLICANT: Norris, Susan R.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)D
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/803,736
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: Identified by Attorney Docket number 04983.0206CPUS01 38-10
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 1582
; SEQ ID NO 500
; LENGTH: 94378
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-803-736-500
```

```
Query Match      80.0%; Score 16.8; DB 6; Length 94378;
Best Local Similarity 90.0%; Pred. No. 2.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY      2 gaagagaccatgaacaactt 21
      ||||||| ||||||| ||
Db      8941 GAAGAGACCATGACCAATT 8922
```

```
RESULT 14
US-09-803-736-1145/c
; Sequence 1145, Application US/09803736
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Levin, Irena M.
; APPLICANT: Norris, Susan R.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)D
; CURRENT APPLICATION NUMBER: US/09/803,736
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/534,859
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: Identified by Attorney Docket number 04983.0206CPUS01 38-10
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 1582
; SEQ ID NO 1145
; LENGTH: 105863
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-803-736-1145
```

```
Query Match      80.0%; Score 16.8; DB 6; Length 105863;
Best Local Similarity 90.0%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY      1 cgaagagaccatgaacaact 20
      ||||||| ||||||| ||
Db      73043 CCAAGAGACCATGAACAACACT 73024
```

```
RESULT 15
US-09-803-110-214/c
; Sequence 214, Application US/09803110
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)D
; CURRENT APPLICATION NUMBER: US/09/803,110
; CURRENT FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/739,449
```

```
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: US 60/168,139
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 214
; LENGTH: 308503
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(308503)
; OTHER INFORMATION: unsure at all n locations
US-09-803-110-214
```

```
Query Match      80.0%; Score 16.8; DB 6; Length 308503;
Best Local Similarity 90.0%; Pred. No. 2.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY      2 gaagagaccatgaacaactt 21
      ||||||| ||||||| ||
Db      262148 GAAGAGACCATGACCAATT 262129
```

Search completed: September 21, 2001, 02:40:59
Job time: 15145 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 00:35:41 : Search time 3479.37 Seconds
(without alignments)
93.357 Million cell updates/sec

Title: US-09-138-735-10
Perfect score: 21
Sequence: 1 acataatgctcctcgtcacac 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_ba3: *
4: gb_in1: *
5: gb_in2: *
6: gb_in3: *
7: gb_on: *
8: gb_ov: *
9: gb_pat1: *
10: gb_pat2: *
11: gb_ph: *
12: gb_p11: *
13: gb_p12: *
14: gb_p13: *
15: gb_p14: *
16: em_ba1: *
17: em_ba2: *
18: em_fun: *
19: em_htgo_hum: *
20: em_htgo_inv: *
21: em_htgo_rod: *
22: em_htg_hum1: *
23: em_htg_hum2: *
24: em_htg_hum3: *
25: em_htg_hum4: *
26: em_htg_hum5: *
27: em_htg_hum6: *
28: em_htg_hum7: *
29: em_htg_hum8: *
30: em_htg_inv1: *
31: em_htg_inv2: *
32: em_htg_other: *
33: em_htg_rod: *
34: em_hum1: *
35: em_hum2: *
36: em_hum3: *
37: em_hum4: *
38: em_hum5: *
39: em_hum6: *
40: em_hum7: *
41: em_in: *
42: em_on: *
43: em_or: *

44: em_ov: *
45: em_pat: *
46: em_ph: *
47: em_p1: *
48: em_ro: *
49: em_sts: *
50: em_sy: *
51: em_un: *
52: em_v1: *
53: gb_sts1: *
54: gb_sts2: *
55: gb_sts3: *
56: gb_sy: *
57: gb_un: *
58: gb_v1: *
59: gb_v12: *
60: gb_v12: *
61: gb_v12: *
62: gb_v12: *
63: gb_v12: *
64: gb_v12: *
65: gb_v12: *
66: gb_v12: *
67: gb_v12: *
68: gb_v12: *
69: gb_v12: *
70: gb_v12: *
71: gb_v12: *
72: gb_v12: *
73: gb_v12: *
74: gb_v12: *
75: gb_v12: *
76: gb_v12: *
77: gb_v12: *
78: gb_v12: *
79: gb_v12: *
80: gb_v12: *
81: gb_v12: *
82: gb_v12: *
83: gb_v12: *
84: gb_v12: *
85: gb_v12: *
86: gb_v12: *
87: gb_v12: *
88: gb_v12: *
89: gb_v12: *
90: gb_v12: *
91: gb_v12: *
92: gb_v12: *
93: gb_v12: *
94: gb_v12: *
95: gb_v12: *
96: gb_v12: *
97: gb_v12: *
98: em_ba3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 21 | 100.0 | 21 | 9 | A48919 Sequence 10 |
| 2 | 21 | 100.0 | 21 | 9 | AR047928 Sequence 10 |
| 3 | 21 | 100.0 | 3402 | 9 | A48910 Sequence 1 |
| 4 | 21 | 100.0 | 3402 | 9 | AR047920 Sequence 1 |
| 5 | 21 | 100.0 | 3402 | 96 | TCU24190 Sequence 1 |
| 6 | 17.8 | 84.8 | 48369 | 63 | AC014450 Drosophila |
| 7 | 17.8 | 84.8 | 153929 | 62 | AC011883 Homo sapi |
| 8 | 17.8 | 84.8 | 213892 | 60 | AC008507 Homo sapi |

| | | | | | | |
|------|------|------|--------|----|-------------|---------------------|
| 9 | 17.8 | 84.8 | 257523 | 5 | AE003812 | AE003812 Drosophila |
| 10 | 17.4 | 82.9 | 163861 | 80 | AL356099 | AL356099 Homo sapi |
| C 11 | 17.4 | 82.9 | 180530 | 65 | AC018613 | AC018613 Homo sapi |
| C 12 | 16.8 | 80.0 | 90824 | 14 | ATT26M18 | AL078606 Arabidops |
| C 13 | 16.8 | 80.0 | 110000 | 84 | LMFLCHR26_3 | Continuation (4 of |
| C 14 | 16.8 | 80.0 | 135161 | 80 | AL357494 | AL357494 Homo sapi |
| C 15 | 16.8 | 80.0 | 143092 | 72 | AC051641 | AC051641 Homo sapi |
| C 16 | 16.8 | 80.0 | 147123 | 70 | AC027030 | AC027030 Homo sapi |
| C 17 | 16.8 | 80.0 | 147945 | 77 | AC084735 | AC084735 Homo sapi |
| C 18 | 16.8 | 80.0 | 149570 | 84 | CNS01RHU | AL162471 Homo sapi |
| C 19 | 16.8 | 80.0 | 164452 | 83 | CNS01RGP | AL159140 Homo sapi |
| C 20 | 16.8 | 80.0 | 190000 | 60 | AC006391 | AC006391 Homo sapi |
| C 21 | 16.8 | 80.0 | 192025 | 88 | AC027129 | AC027129 Homo sapi |
| C 22 | 16.8 | 80.0 | 194140 | 2 | AF242881 | AF242881 Agrobacte |
| C 23 | 16.8 | 80.0 | 197206 | 68 | AC023757 | AC023757 Homo sapi |
| C 24 | 16.8 | 80.0 | 198788 | 13 | ATC8RIV32 | AL161532 Arabidops |
| C 25 | 16.8 | 80.0 | 321708 | 77 | AC087142 | AC087142 Mus muscu |
| C 26 | 16.8 | 80.0 | 338116 | 77 | AC087159 | AC087159 Mus muscu |
| C 27 | 16.4 | 78.1 | 69558 | 78 | AC090544 | AC090544 Homo sapi |
| C 28 | 16.4 | 78.1 | 70250 | 77 | AC090277 | AC090277 Homo sapi |
| C 29 | 16.4 | 78.1 | 70250 | 77 | AC090277 | AC090277 Homo sapi |
| C 30 | 16.4 | 78.1 | 127759 | 75 | AC079021 | AL121975 Human DNA |
| C 31 | 16.4 | 78.1 | 129629 | 93 | HSJ422B11 | AC009885 Homo sapi |
| C 32 | 16.4 | 78.1 | 144163 | 61 | AC009885 | AP002362 Homo sapi |
| C 33 | 16.4 | 78.1 | 147700 | 83 | AP002362 | AC011088 Homo sapi |
| C 34 | 16.4 | 78.1 | 164991 | 87 | AC011088 | AP002457 Homo sapi |
| C 35 | 16.4 | 78.1 | 169168 | 83 | AP002457 | AC006257 Homo sapi |
| C 36 | 16.4 | 78.1 | 175818 | 86 | AC006257 | AP001642 Homo sapi |
| C 37 | 16.4 | 78.1 | 180234 | 82 | AP001642 | AC008506 Homo sapi |
| C 38 | 16.4 | 78.1 | 191557 | 86 | AC008506 | D85291 Pates monke |
| C 39 | 16.2 | 77.1 | 896 | 91 | D85291 | V00672 Chimpanzee |
| C 40 | 16.2 | 77.1 | 896 | 97 | MIRP45 | D45835 Drosophila |
| C 41 | 16.2 | 77.1 | 2278 | 5 | DRORA | AF292113 Aedes aeg |
| C 42 | 16.2 | 77.1 | 3675 | 5 | AF292113 | AX089750 Sequence |
| C 43 | 16.2 | 77.1 | 10628 | 10 | AX089750 | AJ133811 Plasmodiu |
| C 44 | 16.2 | 77.1 | 10628 | 96 | PFA133811 | AC012694 Drosophila |
| C 45 | 16.2 | 77.1 | 55754 | 63 | AC012694 | |

ALIGNMENTS

| | | | | | | |
|---|---|-----|-----|-------------|--|--|
| RESULT 1 | | | | | | |
| LOCUS A48919 | 21 bp | DNA | PAT | 07-MAR-1997 | | |
| DEFINITION Sequence 10 from Patent WO9605312. | | | | | | |
| ACCESSION A48919 | | | | | | |
| VERSION A48919.1 GI:2302578 | | | | | | |
| KEYWORDS | | | | | | |
| SOURCE | | | | | | |
| ORGANISM | | | | | | |
| REFERENCE | | | | | | |
| AUTHORS | 1 (bases 1 to 21) | | | | | |
| TITLE | Paranoh-Baccala, G., Lesenechal, M. and Jolivet, M. | | | | | |
| JOURNAL | NOVEL TRYPAOSOMA CRUZI ANTIGEN, AND GENE CODING THEREFOR; THEIR APPLICATION TO THE DETECTION OF CHAGAS' DISEASE | | | | | |
| COMMENT | Patient: MO 9605312-A 10 22-FEB-1996; | | | | | |
| FEATURES | BIO MERIEUX (FR) | | | | | |
| source | Other publication CA 2173957 960222 | | | | | |
| | Other publication AU 3169195 960307 | | | | | |
| | Other publication FR 2723589 960216. | | | | | |
| | location/Qualifiers | | | | | |
| | 1..21 | | | | | |
| | /organism="unidentified" | | | | | |
| | /db_xref="taxon:32644" | | | | | |
| BASE COUNT | 6 a | 7 c | 3 g | 5 t | | |
| ORIGIN | | | | | | |

Query Match 100.0%; Score 21; DB 9; Length 21;
 Best Local Similarity 100.0%; Pred. No. 8.4;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | | | | |
|--|---|-----------------------|-----|-------------|--|--|
| QY | 1 | acataatgacctgttacac | 21 | | | |
| Db | 1 | ACATAATGCGCTGCTTCACAC | 21 | | | |
| RESULT 2 | | | | | | |
| LOCUS AR047928 | 21 bp | DNA | PAT | 29-SEP-1999 | | |
| DEFINITION Sequence 10 from patent US 5820864. | | | | | | |
| ACCESSION AR047928 | | | | | | |
| VERSION AR047928.1 GI:5970271 | | | | | | |
| KEYWORDS | | | | | | |
| SOURCE | | | | | | |
| ORGANISM | | | | | | |
| REFERENCE | | | | | | |
| AUTHORS | 1 (bases 1 to 21) | | | | | |
| TITLE | Paranoh-Baccala, G., Lesenechal, M. and Jolivet, M. | | | | | |
| JOURNAL | TRYPAOSOMA CRUZI ANTIGEN, gene encoding therefor and methods of detecting and treating chagas disease | | | | | |
| FEATURES | Patient: US 5820864-A 10 13-OCT-1996; | | | | | |
| source | location/Qualifiers | | | | | |
| | 1..21 | | | | | |
| | /organism="unknown" | | | | | |
| BASE COUNT | 6 a | 7 c | 3 g | 5 t | | |
| ORIGIN | | | | | | |

Query Match 100.0%; Score 21; DB 9; Length 21;
 Best Local Similarity 100.0%; Pred. No. 8.4;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | | | | |
|--|---|-----------------------|-------|-------------|--|--|
| QY | 1 | acataatgacctgttacac | 21 | | | |
| Db | 1 | ACATAATGCGCTGCTTCACAC | 21 | | | |
| RESULT 3 | | | | | | |
| LOCUS A48910/c | 3402 bp | DNA | PAT | 07-MAR-1997 | | |
| DEFINITION Sequence 1 from Patent WO9605312. | | | | | | |
| ACCESSION A48910 | | | | | | |
| VERSION A48910.1 GI:2302570 | | | | | | |
| KEYWORDS | | | | | | |
| SOURCE | | | | | | |
| ORGANISM | | | | | | |
| REFERENCE | | | | | | |
| AUTHORS | Trypanosoma cruzi. | | | | | |
| TITLE | Trypanosoma cruzi. | | | | | |
| JOURNAL | Eukaryota; Euzoenozoa; Kinetoplastida; Trypanosomatidae; | | | | | |
| COMMENT | Trypanosoma; Schizotrypanum. | | | | | |
| FEATURES | 1 (bases 1 to 3402) | | | | | |
| source | NOVEL TRYPAOSOMA CRUZI ANTIGEN, AND GENE CODING THEREFOR; THEIR APPLICATION TO THE DETECTION OF CHAGAS' DISEASE | | | | | |
| | Patient: MO 9605312-A 1 22-FEB-1996; | | | | | |
| | BIO MERIEUX (FR) | | | | | |
| | Other publication CA 2173957 960222 | | | | | |
| | Other publication AU 3169195 960307 | | | | | |
| | Other publication FR 2723589 960216. | | | | | |
| | location/Qualifiers | | | | | |
| | 1..3402 | | | | | |
| | /organism="Trypanosoma cruzi" | | | | | |
| | /strain="G" | | | | | |
| | /db_xref="taxon:5693" | | | | | |
| | /dev_stage="EPIMASTIGOTE" | | | | | |
| BASE COUNT | 889 a | 818 c | 958 g | 737 t | | |
| ORIGIN | | | | | | |

Query Match 100.0%; Score 21; DB 9; Length 3402;
 Best Local Similarity 100.0%; Pred. No. 4;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 acataatgacctgttacac 21
 Db 2207 ACATAATGCGCTGCTTCACAC 2187

RESULT 4
LOCUS AR047920/c 3402 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5820864.
ACCESSION AR047920
VERSION AR047920.1 GI:5970263
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3402)
Paranhos-Baccala,G., Lesenechal,M. and Jolivet,M.
Trypanosoma cruzi antigen, gene encoding therefor and methods of
detecting and treating chagas disease
Patent: US 5820864-A 1 13-Oct-1998;
JOURNAL Location/Qualifiers
FEATURES
source 1..3402
BASE COUNT 888 a 821 c 956 g 737 t
ORIGIN
Query Match 100.0%; Score 21; DB 9; Length 3402;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 acataatggcctcggtcacac 21
|||||
Db 2207 ACATATGCGCTCGTTCACAC 2187
|||||
RESULT 5
LOCUS TC024190/c 3402 bp mRNA INV 04-AUG-1997
DEFINITION Trypanosoma cruzi Tc40 antigen (Tc40) mRNA, complete cds.
ACCESSION U24190
VERSION U24190.1 GI:790645
KEYWORDS
SOURCE Trypanosoma cruzi.
ORGANISM Trypanosoma cruzi
Eukaryota; Eulenzozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma; Schizotrypanum.
REFERENCE 1 (bases 1 to 3402)
Lesenechal,M., Duret,L., Cano,M.I., Mortara,R.A., Jolivet,M.,
Canarigo,M.E., da Silveira,J.F. and Paranhos-Baccala,G.
Cloning and characterization of a gene encoding a novel
immunodominant antigen of Trypanosoma cruzi
Mol. Biochem. Parasitol. 87 (2), 193-204 (1997)
97391123
JOURNAL 2 (bases 1 to 3402)
MEDLINE
AUTHORS Lesenechal,M., Franco Da Silveira,J., Mortara,R.A., Duret,L.,
Canarigo,M.E., Jolivet,M. and Paranhos-Baccala,G.
TITLE Direct Submission
JOURNAL Submitted (06-APR-1995) Mylene Lesenechal, BioMerieux, 22 rue Saint
Jean de Dieu, Lyon 69007, France
FEATURES
source 1..3402
location/Qualifiers
/organism="Trypanosoma cruzi"
/strain="G"
/db_xref="taxon:5693"
/dev_stage="epimastigote"
1..35
/gene="Tc40"
/note="spliced leader"
1..3402
misc_RNA
gene
266..3013
/gene="Tc40"
/codon_start=1
/product="Tc40 antigen"
/protein_id="AAC47657.1"
CDS

/db_xref="GI:790646"
translation="MTVTVDLFNHAKPSNNEGRVWSDVATENPEVRAOVLADSOFL
LATYKRRHVLRYVKRSNLKGLTVRAHAKSPPHAYKFNVRNSVAASAKGEFFWYVIL
DETEASNGRPDLARLVKVFVKLDDPVTIPCFSEFFIAESORPLVLVEPQAALID
SSSLIERDVESLEATLQKNCCTLTTLQPVSENSICVSGSGMTPTTEPTMAACT
LNRSTPSWACCCEGEPYKALHLIDATVEENSVLAASCTGVQWLTGVAEPNLRK
FVIDGSIYAMESSERETFAVFDPRKOLAVNNHSPNPFCTTHMFCOVORNGCFNRT
ADGSCVLADMSNRLTIFHLRSRRRQOPOGKTSVYATKPGCVSSGTDASSSHNT
TSAAPASPPVSPYAPAKAAPPAARSAEPHYGSKITIANLYNIGIVTORVYSTG
APATTSTAVSTTTAPQRTSPYGNRPVTVAGLVAAVNSGASAPPTAAKPGEEK
ASAACETSSVAINATRPALHNASLTPQAPDGLAAVVGSGEVSQSLERESVYTN
SRVLKLPDPTIRDOHEQLNLJLEAOMTELQOSRPTPOTOPRDTSSAKSVETVIL
LIADLSRNITRGVYGVNEAIMLHDEVRRAIKNRLOTOKNIKSRIDELAKEST
TOFTLOTOYENIVYKRELATVGSINGSLVYKENSLOKELNSWSSVEMR
MREELCTRESYAKRKATMPDSSLATSLTSFOGRASRPTILATLISMVREOYRGL
YMLMQPSLLIRFLSILIRENNVYSLTLENVETPNVWCVSLQLLEAATEAKE
VYGVVAIDILSERDQIAONGALGSKLTTRAMRAFERQARSSTSRSLQCLNLEKILQ
S"

BASE COUNT 889 a 818 c 958 g 737 t
ORIGIN
Query Match 100.0%; Score 21; DB 96; Length 3402;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 acataatggcctcggtcacac 21
|||||
Db 2207 ACATATGCGCTCGTTCACAC 2187
|||||
RESULT 6
LOCUS AC014450 48369 bp DNA HTG 16-NOV-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
pieces.
ACCESSION AC014450
VERSION AC014450.1 GI:6436885
KEYWORDS HTG: HTGS_PHASE2.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 48369)
Adams,M. and Venter,J.C.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT For further information on this sequence e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source 1..48369
location/Qualifiers
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
BASE COUNT 14352 a 10109 c 9889 g 14019 t
ORIGIN
Query Match 84.8%; Score 17.8; DB 63; Length 48369;
Best Local Similarity 90.5%; Pred. No. 1; 1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 acataatggcctcggtcacac 21
|||||
Db 32318 ACATATGCGCTCGTTCACAC 32338
|||||
RESULT 7

| LOCUS | AC011883/C | DNA | HTG | 01-APR-2000 |
|------------|---|------------|-----|-------------|
| DEFINITION | Homo sapiens clone Rpl1-17N20. WORKING DRAFT SEQUENCE. 13 unordered pieces | | | |
| ACCESSION | AC011883 | | | |
| VERSION | AC011883.4 | GI:7381815 | | |
| KEYWORDS | HTG; HTGS_PHASE1; HTGS_DRAFT. | | | |
| SOURCE | human. | | | |
| ORGANISM | Homo sapiens | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | |
| AUTHORS | 1 (bases 1 to 153929) | | | |
| TITLE | 1 (bases 1 to 153929) | | | |
| JOURNAL | Unpublished | | | |
| REFERENCE | 2 (bases 1 to 153929) | | | |
| AUTHORS | <p> Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baidwin, J., Barna, N., Beckwith, R., Boguski, J., Boucknight, B., Brown, A., Castle, A., Collins, M., Collins, S., Collins, A., Cooke, P., DeRubeis, R., Dewar, K., Dominko, T., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardy, S., Grant, G., Hargis, B., Heaford, A., Horton, L., Howland, J., Johnson, R., Jones, C., Kan, L., Karas, A., Klein, J., Lebeck, J., Liu, C., Locke, K., MacDonald, P., Marquis, N., McMan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tefaye, S., Tirelli, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W., Zimmer, A. and Zody, M. </p> | | | |
| TITLE | Submitted (15-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA | | | |
| JOURNAL | On Apr 1, 2000 this sequence version replaced gi:6535991. | | | |
| COMMENT | All repeats were identified using RepeatMasker: | | | |
| COMMENT | Smit, A.F.A. & Green, P. (1996-1997) | | | |
| COMMENT | http://ftp.genome.washington.edu/RM/RepeatMasker.html | | | |
| COMMENT | ----- Genome Center | | | |
| COMMENT | Center: Whitehead Institute/ MIT Center for Genome Research | | | |
| COMMENT | Center code: W1BR | | | |
| COMMENT | Web site: http://www.seq.wi.mit.edu | | | |
| COMMENT | Contact: sequence_submissions@genome.wi.mit.edu | | | |
| COMMENT | ----- Project Information | | | |
| COMMENT | Center project name: L3671 | | | |
| COMMENT | Center clone name: 17_N_20 | | | |
| COMMENT | ----- Summary Statistics | | | |
| COMMENT | Sequencing vector: M13; M7815; 100% of reads | | | |
| COMMENT | Chemistry: Dye-terminator Big Dye; 100% of reads | | | |
| COMMENT | Assembly program: Phrap; version 0.960731 | | | |
| COMMENT | Consensus quality: 146450 bases at least Q40 | | | |
| COMMENT | Consensus quality: 149447 bases at least Q30 | | | |
| COMMENT | Consensus quality: 150837 bases at least Q20 | | | |
| COMMENT | Insert size: 151000; agarose-fp | | | |
| COMMENT | Insert size: 152729; sum-of-ctrls | | | |
| COMMENT | Quality coverage: 6.0 in Q20 bases; agarose-fp | | | |
| COMMENT | Quality coverage: 5.9 in Q20 bases; sum-of-ctrls | | | |
| COMMENT | ----- | | | |
| COMMENT | * NOTE: This is a 'working draft' sequence. It currently | | | |
| COMMENT | * consists of 13 contigs. The true order of the pieces | | | |
| COMMENT | * is not known and their order in this sequence record is | | | |
| COMMENT | * arbitrary. Gaps between the contigs are represented as | | | |
| COMMENT | * runs of N, but the exact sizes of the gaps are unknown. | | | |
| COMMENT | * This record will be updated with the finished sequence | | | |
| COMMENT | * as soon as it is available and the accession number will | | | |
| COMMENT | * be preserved. | | | |
| COMMENT | * 1 1414: contig of 1414 bp in length | | | |
| COMMENT | * 1415 1514: gap of 100 bp | | | |
| COMMENT | * 1515 3704: contig of 2190 bp in length | | | |
| COMMENT | * 3705 3804: gap of 100 bp | | | |
| COMMENT | * 3805 8775: contig of 4971 bp in length | | | |
| COMMENT | * 8776 8875: gap of 100 bp | | | |
| COMMENT | * 8876 13172: contig of 4297 bp in length | | | |
| COMMENT | * 13173 13272: gap of 100 bp | | | |

| FEATURES | source | 1. | 153929 | Location/Qualifiers |
|----------|-------------------------------------|---|--------|-----------------------------------|
| * | /organism="Homo sapiens" | | | |
| * | /db_xref="taxon:9606" | | | |
| * | /clone="RP11-17N20" | | | |
| * | /clone.lib="RC11-11 Human Male BAC" | | | |
| * | 1..1414 | | | |
| * | misc_feature | /note="assembly-fragment" | | |
| * | misc_feature | 1515..3704 | | |
| * | misc_feature | /note="assembly-fragment" | | |
| * | misc_feature | 3805..8775 | | |
| * | misc_feature | /note="assembly-fragment" | | |
| * | misc_feature | 8876..13172 | | |
| * | misc_feature | /note="assembly-fragment" | | |
| * | misc_feature | 13273..18425 | | |
| * | misc_feature | /note="assembly-fragment" | | |
| * | misc_feature | 18526..27632 | | |
| * | misc_feature | /note="assembly-fragment" | | |
| * | misc_feature | 27733..34151 | | |
| * | misc_feature | /note="assembly-fragment" | | |
| * | misc_feature | 34252..44294 | | |
| * | misc_feature | /note="assembly-fragment" | | |
| * | misc_feature | 44395..54701 | | |
| * | misc_feature | /note="assembly-fragment" | | |
| * | misc_feature | 54802..67345 | | |
| * | misc_feature | /note="assembly-fragment" | | |
| * | misc_feature | 67346..78785 | | |
| * | misc_feature | /note="assembly-fragment" | | |
| * | misc_feature | 78786..103187 | | |
| * | misc_feature | /note="assembly-fragment" | | |
| * | misc_feature | 103188..103287 | | |
| * | misc_feature | /note="assembly-fragment" | | |
| * | misc_feature | 103288..153929 | | |
| * | misc_feature | /note="assembly-fragment" | | |
| * | misc_feature | 67446..78785 | | |
| * | misc_feature | /note="assembly-fragment" | | |
| * | misc_feature | clone_end:r7 | | |
| * | misc_feature | vector_side:right | | |
| * | misc_feature | 78886..103187 | | |
| * | misc_feature | /note="assembly-fragment" | | |
| * | misc_feature | 103288..153929 | | |
| * | misc_feature | /note="assembly-fragment" | | |
| * | misc_feature | clone_end:sp6 | | |
| * | misc_feature | vector_side:right | | |
| * | misc_feature | 42498 a 29651 c 31126 g 49445 t 1209 others | | |
| * | misc_feature | ORIGIN | | |
| * | misc_feature | Query Match | 84.8% | Score 17.8; DB 62; Length 153929; |
| * | misc_feature | Best Local Similarity | 90.5% | Pred. No. 97; |
| * | misc_feature | Matches 19: Conservative 0; Mismatches 2; Indels 0; Gaps 0; | | |
| * | misc_feature | 1 acataatgacctcgttcacac 21 | | |
| * | misc_feature | | | |
| * | misc_feature | Db 64681 ACATAATGACCTCGTTTACAC 64661 | | |
| * | misc_feature | RESULT 8 | | |
| * | misc_feature | AC008507 213892 bp DNA HTG 06-MAY-2000 | | |
| * | misc_feature | LOCUS AC008507 Homo sapiens chromosome 19 clone CTC-448F2, WORKING DRAFT SEQUENCE | | |
| * | misc_feature | DEFINITION 11 ordered pieces. | | |
| * | misc_feature | AC008507 | | |
| * | misc_feature | VERSION AC008507.6 GI:7711281 | | |

KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 213892)
TITLE DOE Joint Genome Institute.
JOURNAL Sequencing of Human Chromosome 19
REFERENCE 2 (bases 1 to 213892)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On May 6, 2000 this sequence version replaced gi:7689761.

-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov.

-----Project Information
Center Project Name: 357334, BC322358
Center Clone Name: CIT-HSPC_448F2

-----Summary Statistics
Consensus quality: 195366 bases at least Q40
Consensus quality: 206085 bases at least Q30
Consensus quality: 210129 bases at least Q20
Estimated insert size: 217290; agarose-ff estimation
Estimated insert size: 212892; sum-of-contigs estimation
Quality coverage: 7.58 in Q20 bases; agarose-ff estimation
Quality coverage: 7.71 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently consists of 11 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
* 1 43108: contig of 43108 bp in length
* 43109 43208: gap of unknown length
* 43209 124836: contig of 81628 bp in length
* 124837 124936: gap of unknown length
* 124937 129045: contig of 4109 bp in length
* 129046 129145: gap of unknown length
* 129146 138110: contig of 8965 bp in length
* 138111 147488: contig of 9278 bp in length
* 147489 147588: gap of unknown length
* 147589 150580: contig of 2892 bp in length
* 150581 154788: gap of unknown length
* 154789 154888: contig of 4208 bp in length
* 154889 157934: gap of unknown length
* 157935 158034: contig of 3046 bp in length
* 158035 160157: gap of unknown length
* 160158 160257: contig of 2123 bp in length
* 160258 165114: gap of unknown length
* 165115 213892: contig of 48678 bp in length.
* 165215 213892: contig of 48678 bp in length.

FEATURES
SOURCE

1. 213892
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CITC-448F2"
/clone_lib="Caltech human BAC library C"
BASE COUNT 54302 a 45251 c 45136 g 68203 t 1000 others
ORIGIN

Query Match 84.8%; Score 17.8; DB 60; Length 213892;
Best Local Similarity 90.5%; Pred. No. 92;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 acataatgacctcgttcacac 21
||||| ||||| ||||| |||||
Db 100316 ACATATGACCTCGTTACAC 100296

RESULT 9
AE003812

LOCUS 257523 bp DNA INV 04-OCT-2000
DEFINITION Drosophila melanogaster genomic scaffold 142000013386047 section 31
of 52 complete sequence.

ACCESSION AE003812 AE002787
VERSION AE003812.2 GI:10727533

KEYWORDS HTG.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster

REFERENCE Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
AUTHORS 1 (bases 1 to 257523)
Adams,M.D., Celinker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galie,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X., Branton,R.C., Rogers,Y.H., Blazey,R.G., Champs,M., Pfeiffer,B.D., Man,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J., Andrews-Planck,K.C., Baldwin,D., Ballew,R.M., Beeson,K.Y., Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y., Benos,P.V., Bernan,B.P., Bhandari,D., Bolshakov,S., Borkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brotlier,P., Burtis,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A., Candia,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de Pablo,J.B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferreira,S., Fleischmann,W., Foster,C., Gabrielian,A.E., Garg,N.S., Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J., Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wel,M.H., Ibegwan,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z., Kienison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C., Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nuskern,D.R., Paclob,J.M., Palazolo,M., Pittman,G.S., Pan,S., Pollard,J., Puri,Y., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Klimos,I., Simpson,M., Skupski,M.P., Smith,T., Spler,E., Spidling,A.C., Stapleton,M., Strong,R., Sun,E., Svitskas,R., Tector,C., Turner,R., Venier,E., Wang,A.H., Wang,X., Wang,Z.Y., Wasserman,D.A., Weinstein,G.M., Weissbach,J., Williams,S.M., Woodde,T., Wortley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F., Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.

TITLE The genome sequence of Drosophila melanogaster
JOURNAL Science 287 (5461), 2185-2195 (2000)

MEDLINE 20196006
REFERENCE 2 (bases 1 to 257523)
AUTHORS Adams,M.D., Celinker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,J.C.
TITLE Direct Submission
JOURNAL Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA

COMMENT On Oct 9, 2000 this sequence version replaced gi:7303102.

FEATURES
SOURCE 1. 257523
/organism="Drosophila melanogaster"

/db_xref="taxon:7227"
/chromosome="2R"
/gene="Cyp317a1"
/product="CT31993"
/db_xref="FLYBASE:FBan0017453"
/db_xref="FLYBASE:FBgn0033982"
/db_xref="GI:7303118"
/map="51D2-51D5"
/db_xref="FLYBASE:FBan0017453"
/db_xref="FLYBASE:FBgn0033982"
/gene="Cyp317a1"
/note="Cyp317a1 gene product"
/codon_start=-1
/db_xref="FLYBASE:FBan0017453"
/db_xref="FLYBASE:FBgn0033982"
/protein_id="AAFS8184.1"
/db_xref="GI:7303118"
/translation="MWIIFLLIGLVGLVLLVLLIARVORDYWRVLDIPHPKRLM
KPIRMNDRKGLVSNOKSDPLSDLYALRGSKMKKDKLPSEGRMSLLTYCLYEE
AEOLLTVNSTLSOPHSTVHIQIMKRVLSLACVFEGLNAEQRTYPLEDEMT
ELALNSHKHGYLNNLMIRPNFCRLRMRTPEKQAEVYFKLLTSIVEQRETSGKQ
KDYLOLLIDVKALEFTHQYEADELCALHQAHAAYVFLKAGYEOTANTLSYLY
ELALHPELOVRYEEVKAIEIRHGDGHITHEGISLSPMGQVNETLMMHPTTYLIR
TLNDVAVPDHKKYILVKEFLIIPTHAIHHDPIYDPEEFKPDMSGDSLOEOT
WRFQVGARSCIGIOFAQQLRLALALLSEYFSLNTRKPLINLEDGILTTLMPGLV
LEPNEERAV"
complement(join(7268..7933,7996..8196,8304..8516,
8578..8824,8929..9357,9866..11244,11307..11746,
16662))
/gene="BCDNA:GH03482"
/product="CT28801"
/db_xref="FLYBASE:FBan0010249"
/db_xref="FLYBASE:FBgn0027596"
complement(join(7268..7933,7996..8196,8304..8516,
8578..8824,8929..9357,9866..11244,11307..11746,
11811..11874))
/gene="BCDNA:GH03482"
/product="CT38587"
/db_xref="FLYBASE:FBan0010249"
/db_xref="FLYBASE:FBgn0027596"
complement(<7268..>16662)
/gene="BCDNA:GH03482"
/note="CG10249"
/map="51D2-51D6"
/db_xref="FLYBASE:FBan0010249"
/db_xref="FLYBASE:FBgn0027596"
complement(join(7766..7933,7996..8196,8304..8516,
8578..8824,8929..9357,9866..11244,11307..11746,
11811..11862))
/gene="BCDNA:GH03482"
/note="BCDNA:GH03482 gene product [alt 2]"
/codon_start=-1
/db_xref="FLYBASE:FBan0010249"
/db_xref="FLYBASE:FBgn0027596"
/protein_id="AAFS8183.1"
/db_xref="GI:7303117"
/translation="MASIQIRIYLRREKSNCSICIGNOSPTGTPRPLASNTLPLRO
RSPPAEYQESYVQATVYANAKSPVEDLESKSPGAPPPPRRHLLATTPGKPTF
KEDRPQENGSSVSSLSQSPVISEGOTTAQALFQIRQMALSLRLKLEQVYKVI
PDMELELNSLRSEKOKLRNEBELMSOROTPSPPSPGLKTASPVQFTPORLSPVSE
SLGARMSSTISASPSVYRVSTQAVVAPATRRVSGHQPPTTRNVQVSEYLY
SKEDLKXVQALJNAOKOLRKLISVGTOMVYPOROKSGVOTPEPRKLKTHVY
TYKPAIRRENTSCKPDYRHGSSDDRLDDPLCKCAKSKSVACGEPTRPEKKRPAO
MPGRNFTSLGENEIVSVAKTIGCOTPVTLVSSGSQTTTAAVRSTGAOVNQOIS
GVQCATQVSRQDTDKGLRLTRSQTKAEVSSVPPPTTRHVASNTDKVEAEPAVY
RTSHSACNTEEVKRDVCGDIYKPHISICANYSACSCKEATQDLARFSKASPTAT
RTGAVARSTASDSRIIPRKHLSSPSVPEKFRONTYTTLATSPSPVERKARLS
SLQEKPPSVSSFOASEPLPESQSFIRPQKDSIDLILGDELLIVREKRVISWSND

ASPAPLMTSSGTTKSESEPEQEPAMSSSDSPPIDVQIAGARKKSTPLKSSQSGS
OYVIVIEOPATKAOHLQADSEPRKRETSQSLKALKIINSLKLGAKIPSSOS
LKSAKATIDHWFNTNSGKSDPHQVEDYIDYESISLPLEYCVMLSSNGNTAMHY
AVSHGNPDVYSITLDSKCDVNMNNAQVSWLVSJLAKOPAHRRVYERLEKMDY
NIRAKKRCOTALMLAVSHGNGDVNMLDAGADININODEGSGTALMCAAEHGVYDIK
HLISHPECDSLVTVDVDSVAFKIAWAGHADVLLIYHEQMLRSKLPNGDVTRSSL
ISPPRHRKQPSQ"
complement(join(7766..7933,7996..8196,8304..8516,
8578..8824,8929..9357,9866..11244,11307..11746,
16017..16614))
/gene="BCDNA:GH03482"
/note="BCDNA:GH03482 gene product [alt 1]"
/codon_start=-1
/db_xref="FLYBASE:FBan0010249"
/db_xref="FLYBASE:FBgn0027596"
/protein_id="AAFS8182.1"
/db_xref="GI:7303116"
/translation="WGISKLVSCHTTPFHRNRISARKCLYIVLLKQTSHESPDS
PHPIGGGSLHMSQSYTTPCYRPRSSSVPRFTYGSIPRSESPFGTASTCSLKG
VESDAGIYHQHRHRAISPPETRAFLHGLADEVSDERLERTSVKRRKSPYAGSL
MDRNNKLSLSINGHGLKELKSHRGNSTMDRHPDVADSCIGNOSPPTGPRPLAR
SNTLPLRQSPPAVOYESYVQATVYANAKSPVEDLESKSPGAPPPPRRHLLAT
TPGKPTPEKDRPOBNSSVSSLSQSPVISEGOTTAQALFOIREQMALSLRLKD
LEQVYKIPMELELNSLRSEKOKLRNEBELMSOROTPSPPSPGLKTASPVQFTPO
RISPVLSLESGAMRRSSISASPSVYRVSTQAVVAPATRRVSGHQPPTTRNG
VQSEYVYSKEELDKVYEQALJNAOKOLRKLISVGTOMVYPOROKDSQVOTTPERP
KLYHGVYVYKPAIRRENTSCKPDYRHGSSDDRLDDPLCKCAKSKSVACGEPTR
EKKVPAPOMGRNFTSLGENEIVSVAKTIGCOTPVTLVSSGSQTTTAAVRSTGA
VNPOOHSGVQCATQVSRQDTDKGLRLTRSQTKAEVSSVPPPTTRHVASNTDK
EABAPKAVRISHSACNTEEVKRDVCGDIYKPHISICANYSACSCKEATQDLARF
SKASPTATRTGAVARSTASDSRIIPRKHLSSPSVPEKFRONTYTTLATSPSPQ
VERKARLSLQEKPPSVSSFOASEPLPESQSFIRPQKDSIDLILGDELLIVREK
VSIWSNDASPAPLMTSSGTTKSESEPEQEPAMSSSDSPPIDVQIAGARKKSTPL
KSSQSGOYVIVIEOPATKAOHLQADSEPRKRETSQSLKALKIINSLKLGAKIPSS
AKPIISOSLSKAKATIDHWFNTNSGKSDPHQVEDYIDYESISLPLEYCVMLSSNG
NCTAMHYAVSHGNPDVYSITLDSKCDVNMNNAQVSWLVSJLAKOPAHRRVYER
RLRKMADVNIIRAKKRCOTALMLAVSHGNGDVNMLDAGADININODEGSGTALMCAE
HGVADYIKHLISHPECDSLVTVDVDSVAFKIAWAGHADVLLIYHEQMLRSKLPAR
GDATRSSLISPPRHRKQPSQ"
complement(join(<35942..36331,36391..36495,36856..38044,
38106..38192,38559..>38663))
/gene="CG10253"
/product="CT28813"
/db_xref="FLYBASE:FBan0010253"
/db_xref="FLYBASE:FBgn0033983"
complement(<35942..>38663)
/gene="CG10253"
/map="51D6-51D7"
/db_xref="FLYBASE:FBan0010253"
/db_xref="FLYBASE:FBgn0033983"
complement(join(35942..36331,36391..36495,36856..38044,
38106..38192,38559..38683))
/gene="CG10253"
/note="CG10253 gene product"
/codon_start=-1
/db_xref="FLYBASE:FBan0010253"
/db_xref="FLYBASE:FBgn0033983"
/protein_id="AAFS8181.1"
/db_xref="GI:7303115"
/translation="MAKRMAVTEAPESAPCEGTALADLSRLKRVESVIRKRRHE
ALKWPGVNDOSQFYCGKDGICFRGKEYPIGGGELPSFTWVKRKLPRADTKQYPO
LPRTPYRPVYNAFYLEKCTQYVDVSAEDILRKHGCTDLIDYLSMHNKRRIID
LVVAPRCHDEVOGLVRLANKHNWLVPEFGGITSVSGITICPOQDSRNICALDTSONMR
LMLNLRNLTFCESGIVGODLERLVLSBELVYGHENEDSYESTLGGVWATRASGMK
NVNGINEDLVAVRVMTVPSGTLRECSAPVSCGFNFHNVILISEGTLGIVTEVVLKV

Query Match 84.8% Score 17.8: DB 5; Length 257523;
Best Local Similarity 90.5% Pred. No. 90;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 acataatgctcgtcacac 21
|||||
Db 49929 ACATGATGCGATCGTTCAAC 49949

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 14 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 1307: contig of 1307 bp in length
1308 1407: gap of 100 bp
1408 2624: contig of 1217 bp in length
2625 2724: gap of 100 bp
2725 4985: contig of 2261 bp in length
4986 5085: gap of 100 bp
5086 8358: contig of 3273 bp in length
8359 8458: gap of 100 bp
8459 12190: contig of 3732 bp in length
12191 12290: gap of 100 bp
12291 18074: contig of 5784 bp in length
18075 18174: gap of 100 bp
18175 24644: contig of 6470 bp in length
24645 24744: gap of 100 bp
24745 30454: contig of 5710 bp in length
30455 30554: gap of 100 bp
30555 38846: contig of 8292 bp in length
38847 38946: gap of 100 bp
38947 52810: contig of 13864 bp in length
52811 52910: gap of 100 bp
52911 75545: contig of 22635 bp in length
75546 75645: gap of 100 bp
75646 103821: contig of 28176 bp in length
103822 103921: gap of 100 bp
103922 136270: contig of 32349 bp in length
136271 136371: gap of 100 bp
136371 180530: contig of 44160 bp in length.
Location/Qualifiers
1. 180530

```

source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8"
 /clone="RP11-555D24"
 /clone_lib="RPC1-11 Human Male BAC".

```

misc_feature 1. 1307
  /note="assembly_fragment"
  1408..2624
  /note="assembly_fragment"
  clone_end:T7
  vector_side:left"
  2725..4985
  /note="assembly_fragment"
  5086..8358
  /note="assembly_fragment"
  8459..12190
  /note="assembly_fragment"
  12291..18074
  /note="assembly_fragment"
  18175..24644
  /note="assembly_fragment"
  24745..30454
  /note="assembly_fragment"
  30555..38846
  /note="assembly_fragment"
  clone_end:SP6
  vector_side:right"
  38947..52810
  /note="assembly_fragment"
  52911..75545
  /note="assembly_fragment"
  75646..103821
  /note="assembly_fragment"
  103922..136270
  /note="assembly_fragment"

```

```

misc_feature 136371..180530
  /note="assembly_fragment"
BASE COUNT 55695 a 33884 c 33497 g 56151 t 1303 others
ORIGIN

```

```

Query Match 82.9% Score 17.4; DB 65; Length 180530;
Best Local Similarity 94.7%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 23179 CATPATGCGCTCGTCAAA 23161
|||||

```

```

RESULT 12
ATT26M18/c
LOCUS ATT26M18 90824 bp DNA PLN 15-MAR-2000
DEFINITION Arabidopsis thaliana DNA chromosome 4, BAC clone T26M18 (SSA
project).
ACCESSION AL078606
VERSION AL078606.1 GI:5002514
KEYWORDS
SOURCE

```

ORGANISM
 thale cress.
 Arabidopsis thaliana

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
 AUTHORS Bevan,M., Volckaert,G., Gyomai,B., Voelt,M., Robben,J.,
 Bancroft,I., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.

JOURNAL
 REFERENCE 2 (bases 1 to 90824)
 AUTHORS Bevan,M., Hildebrandt,H., Braun,M., Holzer,E., Brandt,A.,
 Duesterhoeft,A., Bancroft,I., Mewes,H.W., Lemcke,K. and
 Mayer,K.F.X.

JOURNAL
 REFERENCE 3 (bases 1 to 90824)
 AUTHORS EU Arabidopsis sequencing project.
 TITLE Direct Submission
 JOURNAL Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
 Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
 lemcke@mpi-biochem.mpg.de, mayer@mpi-biochem.mpg.de Project
 Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
 Laboratory, John Innes Centre, Colney Lane, NR4 7UT Norwich, UK,
 E-mail: michael.bevan@bsrc.ac.uk

COMMENT
 Information on performance of analysis and a more detailed
 annotation of this entry and other sequences of chromosomes 3, 4
 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.

FEATURES
 source location/Qualifiers
 1. 90824

```

/organism="Arabidopsis thaliana"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"

```

```

misc_feature 1. 6968
  /note="overlap to BAC T5C23; please refer to this entry
  for analysis and annotation"
  complement(7302..7490)
  /gene="AT4g11800"

```

```

gene /number=1
  7302..12110
  /gene="AT4g11800"
  complement(join(7302..7490,7664..7948,8032..8183,
  8370..8667,8866..8967,9050..9382,9653..9852,9915..10138,
  10220..10449,10560..10647,10762..10860,10991..11106,
  11193..11441,11552..11777,11863..12110))

```

```

CDS /gene="AT4g11800"
  complement(join(7302..7490,7664..7948,8032..8183,
  8370..8667,8866..8967,9050..9382,9653..9852,9915..10138,
  10220..10449,10560..10647,10762..10860,10991..11106,
  11193..11441,11552..11777,11863..12110))
  /gene="AT4g11800"

```


Best Local Similarity 90.0%; Pred. No. 3.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 2 cataatgcctcgttcacac 21
|||||
Db 43650 CATATATGCTCCTTAACTC 43631

RESULT 13
LMFLCHR26_3/c
WPCOMMENT

Sequence split into 7 fragments LOCUS LMFLCHR26 Accession AL160493

| Fragment Name | Begin | End |
|---------------|--------|--------|
| LMFLCHR26_0 | 1 | 110000 |
| LMFLCHR26_1 | 100001 | 210000 |
| LMFLCHR26_2 | 200001 | 310000 |
| LMFLCHR26_3 | 300001 | 410000 |
| LMFLCHR26_4 | 400001 | 510000 |
| LMFLCHR26_5 | 500001 | 610000 |
| LMFLCHR26_6 | 600001 | 677114 |

Continuation (4 of 7) of LMFLCHR26 from base 300001 (AL160493 Leishmania major chromosome)

Query Match 80.0%; Score 16.8; DB 84; Length 110000;
Best Local Similarity 90.0%; Pred. No. 3.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 cataatgcctcgttcacac 21
|||||
Db 13686 CATGATGCTCCTCATTCACAC 13667

RESULT 14
AL357494/c
LOCUS AL357494
DEFINITION Homo sapiens chromosome 9 clone RP11-544D16, *** SEQUENCING IN
PROGRESS ***, 41 unordered pieces.
ACCESSION AL357494
VERSION AL357494.2 GI:9930947
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 135161)
Plumb,B:
Direct Submission
Submitted (20-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerrysanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Aug 27, 2000 this sequence version replaced gi:9214117.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerrysanger.ac.uk
----- Project Information
Center project name: BA544D16
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid: L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 115360 bases at least Q40
Consensus quality: 123354 bases at least Q30
Consensus quality: 127769 bases at least Q20
Insert size: 131161; sum-of-contigs
requests: clonerequest@sanger.ac.uk
Quality coverage: 2.72x in Q20 bases; sum-of-contigs quality
coverage: 1.97x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 41 contigs. The true order of the pieces
* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
2004 2103: contig of 2003 bp in length
2104 4562: contig of 2459 bp in length
4563 4662: gap of 100 bp
4663 7077: contig of 2415 bp in length
7078 7177: gap of 100 bp
7178 11551: contig of 4374 bp in length
11552 11651: gap of 100 bp
11652 15679: contig of 4028 bp in length
15680 15779: gap of 100 bp
15780 18624: contig of 2845 bp in length
18625 18724: gap of 100 bp
18725 20904: contig of 2180 bp in length
20905 21004: gap of 100 bp
21005 23239: contig of 2235 bp in length
23240 23339: gap of 100 bp
23340 25954: contig of 2615 bp in length
25955 26054: gap of 100 bp
26055 29157: contig of 3103 bp in length
29158 29257: gap of 100 bp
29258 34391: contig of 5144 bp in length
34392 34491: gap of 100 bp
34492 37384: contig of 2893 bp in length
37385 37484: gap of 100 bp
37485 42457: contig of 4973 bp in length
42458 42557: gap of 100 bp
42558 46523: contig of 3966 bp in length
46524 46623: gap of 100 bp
46624 50070: contig of 3447 bp in length
50071 50170: gap of 100 bp
50171 52281: contig of 2111 bp in length
52282 52381: gap of 100 bp
52382 55523: contig of 3142 bp in length
55524 55623: gap of 100 bp
55624 59578: contig of 3955 bp in length
59579 59678: gap of 100 bp
59679 62900: contig of 3222 bp in length
62901 63000: gap of 100 bp
63001 68650: contig of 5650 bp in length
68651 68750: gap of 100 bp
68751 71181: contig of 2431 bp in length
71182 71281: gap of 100 bp
71282 73344: contig of 2063 bp in length
73345 73444: gap of 100 bp
73445 75750: contig of 2306 bp in length
75751 75850: gap of 100 bp
75851 79287: contig of 3437 bp in length
79288 79387: gap of 100 bp
79388 83699: contig of 4312 bp in length
83700 83799: gap of 100 bp
83800 86186: contig of 2387 bp in length
86187 86286: gap of 100 bp
86287 89413: contig of 3127 bp in length
89414 89513: gap of 100 bp
89514 91714: contig of 2201 bp in length
91715 91814: gap of 100 bp
91815 94370: contig of 2556 bp in length
94371 94470: gap of 100 bp
94471 96671: contig of 2201 bp in length
96672 96771: gap of 100 bp
96772 99003: contig of 2232 bp in length
99004 99103: gap of 100 bp
99104 101431: contig of 2328 bp in length
101432 101531: gap of 100 bp
101532 105517: contig of 3986 bp in length
105518 105617: gap of 100 bp
105618 108670: contig of 3053 bp in length
108671 108770: gap of 100 bp


```

* 108771 115363: contig of 6593 bp in length
* 115364 115463: gap of 100 bp
* 115464 118107: contig of 2644 bp in length
* 118108 118207: gap of 100 bp
* 118208 123149: contig of 4942 bp in length
* 123150 123249: gap of 100 bp
* 123250 127306: contig of 4057 bp in length
* 127307 127406: gap of 100 bp
* 127407 129514: contig of 2108 bp in length
* 129515 129614: gap of 100 bp
* 129615 132908: contig of 3294 bp in length
* 132909 133008: gap of 100 bp
* 133009 135161: contig of 2153 bp in length.
* location/Qualifiers
1. 135161

```

```

FEATURES
source
1. 135161
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-544D16"
/clone_1fb="RPC1-11.2"
1. 2003
/misc_feature
/feature="assembly_fragment:01220"
fragment_chain:1
2104. 4562
/misc_feature
/feature="assembly_fragment:01140"
fragment_chain:1
4663. 7077
/misc_feature
/feature="assembly_fragment:01572"
fragment_chain:2
7178. 11551
/misc_feature
/feature="assembly_fragment:01331"
fragment_chain:2
11552. 15679
/misc_feature
/feature="assembly_fragment:00070"
15780. 18624
/misc_feature
/feature="assembly_fragment:00073"
18725. 20904
/misc_feature
/feature="assembly_fragment:00121"
21005. 23239
/misc_feature
/feature="assembly_fragment:00140"
23340. 225954
/misc_feature
/feature="assembly_fragment:00226"
26055. 29157
/misc_feature
/feature="assembly_fragment:00247"
29258. 34391
/misc_feature
/feature="assembly_fragment:00314"
34492. 37384
/misc_feature
/feature="assembly_fragment:00347"
37485. 42457
/misc_feature
/feature="assembly_fragment:00375"
42558. 46523
/misc_feature
/feature="assembly_fragment:00409"
46624. 50070
/misc_feature
/feature="assembly_fragment:00506"
50171. 52281
/misc_feature
/feature="assembly_fragment:00536"
52382. 55523
/misc_feature
/feature="assembly_fragment:00553"
55624. 59578
/misc_feature
/feature="assembly_fragment:00630"
59679. 62900
/misc_feature
/feature="assembly_fragment:00660"
63001. 68650
/misc_feature
/feature="assembly_fragment:00805"
68751. 71181
/misc_feature
/feature="assembly_fragment:00856"
71282. 73344
/misc_feature
/feature="assembly_fragment:00859"
73445. 75750
/misc_feature
/feature="assembly_fragment:00939"
75851. 79287
/misc_feature
/feature="assembly_fragment:00965"
79388. 83699

```

```

misc_feature
/feature="assembly_fragment:00966"
83800. 86186
/misc_feature
/feature="assembly_fragment:00982"
86287. 89413
/misc_feature
/feature="assembly_fragment:00992"
89514. 91714
/misc_feature
/feature="assembly_fragment:01041"
91815. 94370
/misc_feature
/feature="assembly_fragment:01097"
94471. 96671
/misc_feature
/feature="assembly_fragment:01159"
96772. 99003
/misc_feature
/feature="assembly_fragment:01181"
99104. 101431
/misc_feature
/feature="assembly_fragment:01240"

```

```

Query Match 80.0%; Score 16.8; DB 80; Length 135161;
Best Local Similarity 90.0%; Pred. No. 3.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 2 cataatggcctgcacac 21
Db 71926 CATGATGCGCTGTCACAC 71907

```

```

RESULT 15
LOCUS AC051641 143092 bp DNA HTG 28-MAY-2000
DEFINITION Homo sapiens chromosome 8 clone RP11-458L9 map 8, LOW-PASS SEQUENCE
ACCESSION AC051641
VERSION AC051641.2 GI:8099851
KEYWORDS HTG: HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 143092)
BIRREN,B., LINTON,L., Nusbaum,C. and Lander,E.
JOURNAL Homo sapiens chromosome 8, clone RP11-458L9
REFERENCE Unpublished
2 (bases 1 to 143092)
BIRREN,B., LINTON,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
ANDERSON,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
BOGUSLAVSKY,L., Boukhalter,B., Brown,A., Burrell,G.,
CAMPIONARIO,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
COLLYMORE,A., Cooke,P., DeRubeis,K., Dewar,K., Diaz,J.S.,
DODGE,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
GALAGAN,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
GRAND-PIERRE,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
HOWLAND,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
KLEIN,J., Larocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
LEVINE,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Margulis,N.,
MCARTHUR,M., McEwan,P., McGuck,A., McKernan,K., McPheeters,R.,
MELDIRIM,T., Meneus,L., Mihova,T., Miranda,C., Mlepa,V., Morrow,J.,
MURPHY,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'NEIL,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
PISANI,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
ROY,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
STANGE-THOMANN,N., Stojanovic,N., Subramanian,A., Talamas,J.,
TESTAYE,S., Theodore,J., Tirrell,A., Travers,M., Trifillo,J.,
VASSILIEV,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
YOUNG,G., Zainoun,J., Zimmer,A. and Zody,M.

```

```

TITLE Direct Submission
JOURNAL Submitted (15-APR-2000) Whitehead Institute/MIT Center for Genome
RESEARCH, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On May 28, 2000 this sequence version replaced gi:7574860.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

```

```

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

```

Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L8717
Center clone name: 458_L₉

NOTE: This record contains 158 individual
sequencing reads that have not been assembled into
contigs. Runs of N are used to separate the reads
and the order in which they appear is completely
arbitrary. Low-pass sequence sampling is useful for
identifying clones that may be gene-rich and allows
overlap relationships among clones to be deduced.
However, it should not be assumed that this clone
will be sequenced to completion. In the event that
the record is updated, the accession number will
be preserved.

1
810 909: contig of 809 bp in length
910 1702: contig of 793 bp in length
1703 1802: gap of 100 bp
1803 2634: contig of 832 bp in length
2635 2734: gap of 100 bp
2735 3544: contig of 810 bp in length
3545 3644: gap of 100 bp
3645 4438: contig of 794 bp in length
4439 4538: gap of 100 bp
4539 5346: contig of 808 bp in length
5347 5446: gap of 100 bp
5447 6240: contig of 794 bp in length
6241 6340: gap of 100 bp
6341 7128: contig of 788 bp in length
7129 7228: gap of 100 bp
7229 8017: contig of 789 bp in length
8018 8117: gap of 100 bp
8118 8924: contig of 807 bp in length
8925 9024: gap of 100 bp
9025 9807: contig of 783 bp in length
9808 9907: gap of 100 bp
9908 10725: contig of 818 bp in length
10726 10825: gap of 100 bp
10826 11629: contig of 804 bp in length
11630 11729: gap of 100 bp
11730 12528: contig of 799 bp in length
12529 12628: gap of 100 bp
12629 13421: contig of 793 bp in length
13422 13521: gap of 100 bp
13522 14323: contig of 802 bp in length
14324 14423: gap of 100 bp
14424 15212: contig of 789 bp in length
15213 15312: gap of 100 bp
15313 16112: contig of 800 bp in length
16113 16212: gap of 100 bp
16213 17003: contig of 791 bp in length
17004 17103: gap of 100 bp
17104 17917: contig of 814 bp in length
17918 18017: gap of 100 bp
18018 18831: contig of 814 bp in length
18832 18931: gap of 100 bp
18932 19726: contig of 795 bp in length
19727 19826: gap of 100 bp
19827 20585: contig of 759 bp in length
20586 20685: gap of 100 bp
20686 21487: contig of 802 bp in length
21488 21587: gap of 100 bp
21588 22392: contig of 805 bp in length
22393 22492: gap of 100 bp
22493 23294: contig of 802 bp in length
23295 23394: gap of 100 bp
23395 24220: contig of 826 bp in length
24221 24320: gap of 100 bp
24321 25113: contig of 793 bp in length
25114 25213: gap of 100 bp

25214 26008: contig of 795 bp in length
26009 26108: gap of 100 bp
26109 26927: contig of 819 bp in length
26928 27027: gap of 100 bp
27028 27810: contig of 783 bp in length
27811 27910: gap of 100 bp
27911 28724: contig of 814 bp in length
28725 28824: gap of 100 bp
28825 29646: contig of 822 bp in length
29647 29746: gap of 100 bp
29747 30569: contig of 823 bp in length
30570 30669: gap of 100 bp
30670 31460: contig of 791 bp in length
31461 31560: gap of 100 bp
31561 32335: contig of 775 bp in length
32336 32435: gap of 100 bp
32436 33231: contig of 796 bp in length
33232 33331: gap of 100 bp
33332 34131: contig of 800 bp in length
34132 34231: gap of 100 bp
34232 35017: contig of 786 bp in length
35018 35117: gap of 100 bp
35118 35903: contig of 786 bp in length
35904 36003: gap of 100 bp
36004 36812: contig of 809 bp in length
36813 36912: gap of 100 bp
36913 37735: contig of 823 bp in length
37736 37835: gap of 100 bp
37836 38635: contig of 800 bp in length
38636 38735: gap of 100 bp
38736 39551: contig of 816 bp in length
39552 39651: gap of 100 bp
39652 40476: contig of 825 bp in length
40477 40576: gap of 100 bp
40577 41389: contig of 813 bp in length
41390 41489: gap of 100 bp
41490 42301: contig of 812 bp in length
42302 42401: gap of 100 bp
42402 43164: contig of 763 bp in length
43165 43264: gap of 100 bp
43265 43942: contig of 678 bp in length
43943 44042: gap of 100 bp
44043 44820: contig of 778 bp in length
44821 44920: gap of 100 bp
44921 45726: contig of 806 bp in length
45727 45826: gap of 100 bp
45827 46616: contig of 790 bp in length
46617 46716: gap of 100 bp
46717 47549: contig of 833 bp in length
47550 47649: gap of 100 bp
47650 48465: contig of 816 bp in length
48466 48565: gap of 100 bp
48566 49387: contig of 822 bp in length
49388 49487: gap of 100 bp
49488 50278: contig of 791 bp in length
50279 50378: gap of 100 bp
50379 51135: contig of 757 bp in length
51136 51235: gap of 100 bp
51236 52012: contig of 777 bp in length
52013 52112: gap of 100 bp
52113 52913: contig of 801 bp in length
52914 53013: gap of 100 bp
53014 53812: contig of 799 bp in length
53813 53912: gap of 100 bp
53913 54695: contig of 783 bp in length
54696 54795: gap of 100 bp
54796 55581: contig of 786 bp in length
55582 55681: gap of 100 bp
55682 56480: contig of 799 bp in length
56481 56580: gap of 100 bp
56581 57410: contig of 830 bp in length
57411 57510: gap of 100 bp
57511 58337: contig of 827 bp in length

```

* 58338 58437: gap of 100 bp
* 58438 59267: contig of 830 bp in length
* 59268 59367: gap of 100 bp
* 59368 60161: contig of 794 bp in length
* 60162 60261: gap of 100 bp
* 60262 61046: contig of 785 bp in length
* 61047 61146: gap of 100 bp
* 61147 61938: contig of 792 bp in length

Query Match      80.0%; Score 16.8; DB 72; Length 143092;
Best Local Similarity 90.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 cataatggcctcggttcacac 21
   |||||||  |||||||
Db 2303 CAGAAATGCCCTTGTTACAC 2322

```

Search completed: September 21, 2001, 00:36:20
 Job time: 29635 sec

THIS PAGE BLANK (USPTO)

CC The primers AAT27311-5 were used to PCR amplify the sequence encoding a
CC novel isolated antigenic protein from Trypanosoma cruzi epimastigotes,
CC designated PTC100t (AAT27310). The primers, derived from the sequences
CC of a 594 and 1041 bp fragment of PTC100t, amplified the gene as 3
CC fragments. This primer corresponds to nucleotides 2187-2207 of the PTC100t
CC sequence, derived from the 1041 bp fragment. The 594 bp fragment was
CC isolated from a T. cruzi genomic expression library in lambda gtl1, using
CC a mixture of sera from patients with Chagas disease. It corresponds to
CC nucleotides 1232-1825 of PTC100t. The 1041 bp fragment was isolated from
CC a lambda gtl10 library using the 594 bp fragment as a probe. The protein,
CC or antibodies raised against it, can be used in the detection and
CC monitoring of T. cruzi infection i.e. Chagas disease.
XX
SQ Sequence 21 BP; 6 A; 7 C; 3 G; 5 T; 0 other;

Query Match 100.0%; Score 21; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.061; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 acataatgctcgttcacac 21
|||||
Db 1 acataatgctcgttcacac 21

RESULT 2
AAx84096
ID AAx84096 standard; DNA; 21 BP.
XX
AC AAx84096;
XX
DT 27-AUG-1999 (first entry)
XX
DE PCR primer for T. cruzi PTC40 coding sequence.
XX
Ptc40: Tc40; infection; diagnosis; immune complex; antigenic determinant;
KM therapy; antibody; PCR primer; ss.
XX
OS Synthetic.
XX Trypanosoma cruzi.
XX
PN MO92929867-A1.
XX
PD 17-JUN-1999.
XX
PF 10-DEC-1998; 98WO-IB01987.
XX
PR 10-DEC-1997; 97US-0988242.
XX
PA (INMR) BIO MERIEUX.
XX
PI Jolivet M, Lesenechal M, Mandrand B, Paranhos-Baccala G;
XX WPI; 1999-394978/33.
XX
DR WPI; 1999-394978/33.
XX
PT New Trypanosoma cruzi antigen
XX
PS Disclosure; Page 21; 65pp; English.
XX
XX This sequence is a PCR primer for DNA encoding the Trypanosoma cruzi
CC PTC40 protein of the invention. The PTC40 antigenic determinant is
CC useful as a reagent for detection and/or monitoring of Trypanosoma cruzi
CC infection from samples including blood serum or plasma, urine, saliva, or
CC tears, by contacting with the sample and detecting an immune complex. The
CC PTC40 antigenic determinant, the vector, expression cassette, cell or
CC antibody are useful for treatment or prevention (vaccine) of a
CC Trypanosoma cruzi infection in a man or animal. Current Trypanosoma cruzi
CC antigens are obtained from protein fractions of the noninfectious stage
CC of the parasite, and these do not allow sufficient production of antigens
CC for use in reliable serological diagnostic tests. The strain to strain
CC polymorphism reduces reliability of the tests.
XX
SQ Sequence 21 BP; 6 A; 7 C; 3 G; 5 T; 0 other;

Query Match 100.0%; Score 21; DB 20; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.061; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 acataatgctcgttcacac 21
|||||
Db 1 acataatgctcgttcacac 21

RESULT 3
AAT27310/c
ID AAT27310 standard; cDNA; 3402 BP.
XX
AC AAT27310;
XX
DT 26-NOV-1996 (first entry)
XX
DE Trypanosoma cruzi epimastigotic PTC100t antigen gene.
XX
KM Antigen; Trypanosoma cruzi; epimastigote; serum; Chagas disease; probe;
KM Primer; PCR; polymerase chain reaction; amplification; antibody; ds.
XX
OS Trypanosoma cruzi.
XX
XX Key Location/Qualifiers
FT CDS 266..3013
FT /*tag- a
FT /product- PTC100t epimastigotic antigen
XX
PN FR2723589-A1.
XX
PD 16-FEB-1996.
XX
PF 12-AUG-1994; 94FR-0010132.
XX
PR 12-AUG-1994; 94FR-0010132.
XX
PA (INMR) BIO MERIEUX.
XX
PI Jolivet M, Lesenechal M, Paranhos-Baccala G;
XX WPI; 1996-190287/20.
XX
DR P-PSDB; AAR91615.
XX
PT New nucleic acid encoding Trypanosoma cruzi epimastigotic antigen -
XX useful for diagnosis, monitoring and therapy of Chagas disease
XX
PS Claim 1; Page 24-26; 55pp; French.
XX
XX This is the nucleotide sequence encoding a novel isolated antigenic
CC protein from Trypanosoma cruzi epimastigotes, designated PTC100t.
CC The clone Tc50 was isolated from a T. cruzi genomic expression library in
CC lambda gtl1, using a mixture of sera from patients with Chagas disease.
CC Clone Tc50 contained an 594 bp insert corresp. to nucleotides 1232-1825
CC of this sequence. The Tc50 sequence was subsequently used to probe a
CC Southern blot of restriction enzyme digested T. cruzi DNA and also screen
CC a lambda gtl10 library to isolate a 1041 bp EcoRI fragment corresp. to
CC nucleotides 1403-2443 of PTC100t. Primers (AAT27311-5) were synthesised
CC based on the sequences of the 594 and 1041 bp fragments and used to
CC amplify the PTC100t clone as 3 fragments from cDNA derived from mRNA
CC purified from T. cruzi epimastigotes. The protein or antibodies raised
CC against it can be used in the detection and monitoring of T. cruzi
CC infection i.e. Chagas disease.
XX
SQ Sequence 3402 BP; 889 A; 818 C; 958 G; 737 T; 0 other;

Query Match 100.0%; Score 21; DB 17; Length 3402;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 acataatggcctcgcttcacac 21
 |||||||
 DB 2207 ACATATATGGCCTCGTTCAACAC 2187

RESULT 4

AAx84092/C
 ID AAx84092 standard; cDNA; 3402 BP.

AC AAx84092;

DT 27-AUG-1999 (first entry)

DE T. cruzi PTC40 coding sequence.

KW PTC40: Tc40; infection; diagnosis; immune complex; antigenic determinant; therapy; antibody; ds.

OS Trypanosoma cruzi.

PN WO929867-A1.

PD 17-JUN-1999.

PF 10-DEC-1998; 98WO-IB01987.

PR 10-DEC-1997; 97US-0988242.

PA (INMR) BIO MERIEUX.

PI Jolivet M, Lesenechal M, Mandrand B, Paranhos-Baccala G;

DR WPI: 1999-394978/33.

PT P-PSDB; AAY22124.

PS New Trypanosoma cruzi antigen

PS Claim 1; Page 52-56; 65pp; English.

CC This sequence encodes the Trypanosoma cruzi PTC40 protein of the invention, and is designated Tc40. The PTC40 antigenic determinant is useful as a reagent for detection and/or monitoring of Trypanosoma cruzi infection from samples including blood serum or plasma, urine, saliva, or tears, by contacting with the sample and detecting an immune complex. The PTC40 antigenic determinant, the vector, expression cassette, cell or antibody are useful for treatment or prevention (vaccine) of a Trypanosoma cruzi infection in a man or animal. Current Trypanosoma cruzi antigens are obtained from protein fractions of the noninfectious stage of the parasite, and these do not allow sufficient production of antigens for use in reliable serological diagnostic tests. The strain to strain polymorphism reduces reliability of the tests.

CC Sequence 3402 BP; 888 A; 821 C; 956 G; 737 T; 0 other;

Query Match 100.0%; Score 21; DB 20; Length 3402;
 Best Local Similarity 100.0%; Pred. No. 0.13;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 acataatggcctcgcttcacac 21
 |||||||
 DB 2207 ACATATATGGCCTCGTTCAACAC 2187

RESULT 5

AAf57301
 ID AAf57301 standard; DNA; 10628 BP.

AC AAf57301;

DT 29-MAY-2001 (first entry)

DE P. falciparum FCR3.varCSA protein encoding DNA.

XX FCR3.varCSA protein; chondroitin sulfate A; CSA; var gene; PFEPM1;
 KW erythrocyte membrane protein 1; parasitized red blood cell; PRBC;
 KW malaria; protozoasidae; ds.

OS Plasmodium falciparum.

Key Location/Qualifiers
 FT CDS 1..10626
 /tag= a
 /note= "the stop codon is not indicated"

WO200116326-A2.

PD 08-MAR-2001.

PF 01-SEP-2000; 2000WO-US24195.

PR 01-SEP-1999; 99US-0152023.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C;

PT Gysin J, Pouvelle B, Fujii N, Smith J;

DR WPI: 2001-235109/24.

PT P-PSDB; AAB62142.

PS Novel FCR3.varCSA protein, useful for modulating parasitized red blood cell binding, sequestration and onset of maternal malaria -

PS Claim 3; Page 60-63; 78pp; English.

CC The invention relates to a P. falciparum FCR3.varCSA protein, that is capable of binding to chondroitin sulfate A (CSA). The var gene and the corresponding P.falciparum erythrocyte membrane protein 1 (PFEPM1) modulate adhesion of parasitized red blood cell (PRBC) to CSA. The CC protein and the encoding gene are useful for treating and preventing CC maternal malaria in a patient afflicted at a risk for contracting CC maternal malaria or in a patient afflicted with maternal malaria. The CC present sequence represents the DNA encoding the P. falciparum FCR3.varCSA protein.

CC Sequence 10628 BP; 4341 A; 1375 C; 2094 G; 2818 T; 0 other;

Query Match 77.1%; Score 16.2; DB 22; Length 10628;
 Best Local Similarity 85.7%; Pred. No. 47;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 acataatggcctcgcttcacac 21
 ||| |||||
 DB 9358 acaaatgctccctcgcttcacac 9378

RESULT 6
 AA059461
 ID AA059461 standard; cDNA; 296 BP.

AC AA059461;

DT 16-MAR-1994 (first entry)

DE Human brain Expressed Sequence Tag EST01451.

KW Gene transcription product; genetic markers; tagging; in vivo;
 KW transcription; mapping; locations; chromosomes; chromosomal; ss.

OS Homo sapiens.

PN WO9316178-A.

PD 19-AUG-1993.

```

XX 12-FEB-1993; 93WO-US01294.
PF
XX
XX 12-FEB-1992; 92US-0837195.
PR
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.
PI Adams MD, Moreno RF, Venter CJ;
XX
XX WPI; 1993-272882/34.
DR
XX Enriched oligonucleotides and corresp. sequences - used as
PT markers for human genes transcribed in-vivo, facilitate tagging
PT of most human genes
PS
XX Example 4; Page 178; 500pp; English.
XX
XX The Expressed Sequence Tag was isolated from a human brain cDNA
CC library as part of a large set of ESTs which can be used as markers
CC for human genes transcribed in vivo. They can be used to facilitate
CC tagging of most human genes, for mapping locations of expressed genes
CC on chromosomes, for individual or forensic identification, for mapping
CC locations of disease-associated genes, for identification of tissue
CC type, and for prep. of antisense sequences, probes and constructs.
CC EST01451 has a "poor" coding probability as evaluated using the
CC coding-region prediction program CRM. See also AA059041-061440.
CC
XX
XX Sequence 296 BP; 63 A; 72 C; 80 G; 78 T; 3 other:
SQ
Query Match 75.2%; Score 15.8; DB 14; Length 296;
Best Local Similarity 89.5%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 ataatggcctcgtcacac 21
170 agaattgctcgtcacac 188
Db
RESULT 7
AA050584
ID AAX60584 standard; DNA; 1354 BP.
XX
XX AAX60584;
AC
XX
XX 27-JUL-1999 (first entry)
DT
XX
XX Human secreted protein encoding DNA (clone col000_1).
DE
XX
XX Secreted protein; human; tissue marker; genetic disease; gene therapy;
KM veterinary medicine; cell proliferation; immunosuppressant; infection;
KM immunosuppressant; autoimmune disease; organ rejection; tumour; anaemia;
KM haemolysis; wound healing; fertility control; chemotaxis; analgesic;
KM thrombolytic; haemophilia; infarction; antimicrobial agent; cancer; ss.
XX
XX Homo sapiens.
OS
XX
XX WO924469-A1.
PN
XX
XX 20-MAY-1999.
PD
XX
XX 06-NOV-1998; 98WO-US23829.
PF
XX
XX 04-NOV-1998; 98US-0185936.
PR
XX 07-NOV-1997; 97US-0965789.
PR
XX
XX (GEMV ) GENETICS INST INC.
PA
XX
XX Agostino MJ, Evans C, Jacobs K, Lavallie ER, McCoy JM;
PI Werberg D, Racie LA, Treacy M;
XX
XX WPI; 1999-327362/27.
DR
XX
XX P-PSDB; AAY16784.

```

```

XX Nucleic acid encoding secreted human proteins
PT
XX
XX Claim 28; Page 97; 107pp; English.
PS
XX
XX The invention provides polynucleotides (AAX60579-X60687) encoding
CC specific secreted human proteins (AAY16779-Y16787). The nucleic acid
CC sequences are deposited under the accession number ATCC 98580. The
CC polynucleotides are used as tissue markers, chromosomal tags, for
CC diagnosis of genetic diseases, to generate anti-protein or anti-DNA
CC antibodies, also as nutritional sources and supplements and in gene
CC therapy. The secreted proteins are useful therapeutically, in human or
CC veterinary medicine, e.g. for modulating cell proliferation or
CC differentiation, as immunostimulants or immunosuppressants (for treating
CC infections, autoimmune disease, organ rejection, or to induce tumour
CC immunity), as regulators of haematopoiesis (e.g. for treating anaemia or
CC in conjunction with tumour therapy), to stimulate growth of tissue for
CC wound healing, as fertility control agents, for regulating chemotaxis or
CC chemokines (e.g. for directing cells to tumours or sites of infection), as
CC haemostatic and thrombolytic agents (e.g. in treatment of haemophilia or
CC infarctions), as antimicrobial agents, for modifying biorythms,
CC appetite, or metabolism, as analgesics and many other uses. The proteins
CC are also used to raise antibodies, used as diagnostic immunoassay
CC reagents also (when neutralizing) for treating e.g. cancer.
CC
XX
XX Sequence 1354 BP; 344 A; 369 C; 287 G; 350 T; 4 other:
SQ
Query Match 75.2%; Score 15.8; DB 20; Length 1354;
Best Local Similarity 89.5%; Pred. No. 57;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 ataatggcctcgtcacac 21
29 agaattgctcgtcacac 47
Db
RESULT 8
AA051427/c
ID AA051427 standard; cDNA to mRNA; 2344 BP.
XX
XX AA051427;
AC
XX
XX 20-MAY-1994 (first entry)
DT
XX
XX Human FACC cDNA clone #2.
DE
XX
XX Fanconi Anemia Group C; FACC; complementing cDNA; variant; diagnosis;
KM open reading frame; Fanconi anemia; gene therapy; ss.
XX
XX Homo sapiens.
OS
XX
XX Key location/Qualifiers
FH CDS 256..1935
FT CDS /tag- a
FT /product- Human FACC
FT misc-difference 1894..1896
FT /tag- b
FT /note- "Apparent duplication of CTG codon, causing the
FT polyA_signal 2325..2330 protein encoded to be one amino acid longer
FT /tag- c than the protein given in ARA44139"
FT
XX
XX WO9322435-A.
PN
XX
XX 11-NOV-1993.
PD
XX
XX 27-APR-1993; 93WO-CA00178.
PF
XX
XX 29-APR-1992; 92US-0876285.
PR
XX 21-JUL-1992; 92US-0918313.
PR
XX 15-JAN-1993; 93US-0003963.
PR

```


XX (HOSP-) HOSPITAL FOR SICK CHILDREN.
PA (UNME-) UNITED MEDICAL & DENTAL SCHOOL GUYS.
XX
XX
PI Buchwald M, Mathew CG, Strathdee CA, Weyrick R;
XX WPI. 1993-368794/46.
DR P-PSDB; AAR44139.
XX
XX
PT Human cDNA which complements Fanconi Anaemia gp. C - used to
PT develop prods. for use in diagnosis, study and therapy of Fanconi
PT Anaemia
XX
XX
PS Claim 1; Page 101-04; 137pp; English.
XX
XX The sequences given in AA051426-28 represent cDNA variants from the
CC Fanconi Anaemia group C Complementing (FACC) cDNA. These three cDNA
CC molecules are cellular variants of a single cDNA transcribed from the
CC same gene. The three cDNAs each contain an identical open reading
CC frame encoding the FACC protein. The FACC protein may be used for
CC the diagnosis and study of Fanconi anemia.
XX
XX Sequence 2344 BP; 597 A; 594 C; 567 G; 586 T; 0 other;

| | | | | | |
|---------|-----------------------|--------|---------------|--------|--------------|
| | Query Match | 75.2% | Score 15.8; | DB 14; | Length 2344; |
| | Best Local Similarity | 89.5%; | Pred. NO. 61; | | |
| Matches | 17; Conservative | 0; | Mismatches | 2; | Indels 0; |
| Gaps | 0; | | | | |
| Qy | 2 cataatgctcgttcaca | 20 | | | |
| | | | | | |
| db | 939 CAAATGGCTCGTTTACA | 921 | | | |

| | | | |
|----|-------------------------------|--|------------|
| XX | AA051428/c | 9 | |
| XX | AA051428 standard; | CDNA to mRNA; | 3150 BP. |
| XX | AA051428; | | |
| XX | 20-MAY-1994 | (first entry) | |
| XX | Human FACC | CDNA clone #3. | |
| XX | Fanconi Anemia Group C; FACC; | complementing CDNA; variant; | diagnosis; |
| XX | open reading frame; | Fanconi anemia; gene therapy; | ss. |
| XX | Homo sapiens. | | |
| XX | Key | Location/Qualifiers | |
| XX | CDS | 256..1935 | |
| XX | FT | /*tag= a | |
| XX | FT | /product= Human FACC | |
| XX | FT | misc_difference | 1894..1896 |
| XX | FT | /*tag= b | |
| XX | FT | /note= "Apparent duplication of CUG codon, causing the | |
| XX | FT | protein encoded to be one amino acid longer | |
| XX | FT | than the protein given in AA04139" | |
| XX | FT | polyA_signal | 2325..2330 |
| XX | FT | /*tag= c | |
| XX | FT | polyA_signal | 3127..3132 |
| XX | FT | /*tag= d | |
| XX | PN | W09322435-A. | |
| XX | XX | | |
| XX | 11-NOV-1993. | | |
| XX | XX | | |
| XX | 27-APR-1993; | 93WO-CA00178. | |
| XX | XX | | |
| XX | 29-APR-1992; | 92US-0876285. | |
| XX | 21-JUL-1992; | 92US-0918313. | |
| XX | 15-JAN-1993; | 93US-0003963. | |
| XX | XX | | |

PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
PA (UNME-) UNITED MEDICAL & DENTAL SCHOOL GUYS.
XX
XX
PI Buchwald M, Mathew CG, Strathdee CA, Weyrick R;
XX
DR WPI: 1993-368794/46.
DR P-PSDB: AAR44139.
XX
XX Human cDNA which complements Fanconi Anaemia gp. C - used to
PT develop prods. for use in diagnosis, study and therapy of Fanconi
PT Anaemia
XX
PS Claim 1, Page 104-07; 137pp; English.
XX
XX The sequences given in AA051426-28 represent cDNA variants from the
CC Fanconi Anaemia Group C Complementing (FACC) cDNA. These three cDNA
CC molecules are cellular variants of a single cDNA transcribed from the
CC same gene. The three cDNAs each contain an identical open reading
CC frame encoding the FACC protein. The FACC protein may be used for
CC the diagnosis and study of Fanconi anaemia.
XX
XX Sequence 3150 BP; 801 A; 772 C; 771 G; 806 T; 0 other;

| | | | | | |
|----|--------------------------|--------|---------------|-----------|--------------|
| | Query Match | 75.2%; | Score 15.8; | DB 14; | Length 3150; |
| | Best Local Similarity | 89.5%; | Pred. No. 64; | | |
| | Matches 17; Conservative | 0; | Mismatches 2; | Indels 0; | Gaps 0; |
| OY | 2 cataatgctcgttcaca | 20 | | | |
| | | | | | |
| db | 939 CAAATGGCCCTCGTTTACA | 921 | | | |

| XX | RESULT | 10 |
|----|--|-----------------------------------|
| XX | AA051426/c | |
| XX | ID | AA051426 standard; cDNA; 4488 BP. |
| XX | AA051426; | |
| XX | 20-MAY-1994 | (first entry) |
| XX | Human FACC cDNA clone #1. | |
| XX | Panconi Anemia Group C; FACC; Complementing cDNA; variant. | |
| XX | open reading frame; diagnosis; Fanconi anemia; ss. | |
| XX | Homo sapiens. | |
| XX | Key | Location/Qualifiers |
| XX | CDS | 174..1850 |
| XX | | /*tag= a |
| XX | | /product= Human FACC |
| XX | polyA_signal | 2240..2245 |
| XX | | /*tag= b |
| XX | polyA_signal | 3042..3047 |
| XX | | /*tag= c |
| XX | repeat_unit | 3163..3175 |
| XX | | /*tag= d |
| XX | | /rpt_type= OTHER |
| XX | | /note= "palindrome" |
| XX | repeat_unit | 3289..3322 |
| XX | | /*tag= e |
| XX | | /rpt_type= TANDEM |
| XX | repeat_unit | 3323..4455 |
| XX | | /*tag= f |
| XX | | /rpt_type= TANDEM |
| XX | W09322435-A. | |
| XX | 11-NOV-1993. | |
| XX | 27-APR-1993; | 93WO-CA00178. |
| XX | | |

PR 29-APR-1992; 9205-0876285.
 PR 21-JUL-1992; 9205-0918313.
 PR 15-JAN-1993; 9305-0003963.
 XX
 PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
 PA (UNME-) UNITED MEDICAL & DENTAL SCHOOL GUYS.
 PI Buchwald M, Mathew CG, Strathdee CA, Wevrick R;
 XX WPI; 1993-368794/46.
 DR P-PSDB; AAR44139.
 XX
 PT Human cDNA which complements Fanconi Anaemia gp. C - used to
 PT develop prods. for use in diagnosis, study and therapy of Fanconi
 PT Anaemia
 PS Claim 1: Page 97-101; 137pp; English.
 XX
 CC The sequences given in AA051426-28 represent cDNA variants from the
 CC Fanconi Anaemia Group C Complementing (FACC) cDNA. These three cDNA
 CC molecules are cellular variants of a single cDNA transcribed from the
 CC same gene. The three cDNAs each contain an identical open reading
 CC frame encoding the FACC protein. The FACC protein may be used for
 CC the diagnosis and study of Fanconi anemia.
 CC
 SQ Sequence 4488 BP; 1052 A; 1092 C; 1168 G; 1176 T; 0 other;
 SQ
 Query Match 75.2%; Score 15.8; DB 14; Length 4488;
 Best Local Similarity 89.5%; Pred. No. 67;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 2 cataatgacctgtcaca 20
 II |||||
 DB 857 CAAATGCGCTCGTTTACA 839
 DB
 RESULT 11
 AAV33945/C
 ID AAV33945 standard; cDNA; 4567 BP.
 AC AAV33945;
 XX
 DT 15-FEB-1999 (first entry)
 XX
 DE Fanconi anaemia complementation group C (FAC) cDNA.
 XX
 KW Fanconi anaemia complementation group C; FAC; apoptosis;
 KW haematopoiesis; bone marrow; chemotherapy; gene therapy; human; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH FT 256..1929
 FT CDS /*lag- a
 FT
 FT
 PN W09851792-A1.
 XX
 PD 19-NOV-1998.
 XX
 PE 15-MAY-1998; 98WO-US09975.
 XX
 PR 15-MAY-1997; 97US-0046546.
 XX
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL.
 PI Yousoufian H;
 XX
 DR WPI; 1999-009774/01.
 DR P-PSDB; AAW68546.
 XX
 PT New conjugate of Fanconi anaemia molecule and peptide selective for
 PT haematopoietic precursor cells - inhibits apoptosis of these cells,

PT for treating Fanconi anaemia and patients undergoing high-dose
 PT chemotherapy for cancer
 XX
 PS Claim 6: Page 40-45; 72pp; English.
 XX
 CC This cDNA clone includes a coding region for human Fanconi anaemia
 CC complementation group C (FAC, see AAW68546), a protein that modulates
 CC apoptosis in haematopoietic progenitor cells (HPC). The invention
 CC provides conjugates, including fusion proteins, comprising FAC
 CC and a targeting molecule which binds to a cell surface protein of
 CC the HPC and is internalised. Such targeting molecules include
 CC interleukin-3 (see AAW68547) and antibodies which recognise CD33 (see
 CC AAW68548-49). The conjugate, or a nucleic acid encoding it, can be
 CC used to deliver FAC to an HPC, specifically to inhibit apoptosis,
 CC particularly in patients exposed to high doses of chemotherapy for
 CC treatment of non-myeloid cancers, also to treat Fanconi anaemia (by
 CC complementation of the genetic defect). Treatment of HPC is done
 CC in vitro, ex vivo (e.g. for recombinant production of conjugate in
 CC cell cultures) or in vivo. Treatment with FAC may eliminate the
 CC need for extensive bone marrow transplants to restore
 CC haematopoiesis after chemotherapy.
 CC
 SQ Sequence 4567 BP; 1080 A; 1135 C; 1177 G; 1175 T; 0 other;
 SQ
 Query Match 75.2%; Score 15.8; DB 20; Length 4567;
 Best Local Similarity 89.5%; Pred. No. 68;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 2 cataatgacctgtcaca 20
 II |||||
 DB 939 CAAATGCGCTCGTTTACA 921
 DB
 RESULT 12
 AAT39622/C
 ID AAT39622 standard; cDNA; 3417 BP.
 AC AAT39622;
 XX
 DT 30-DEC-1996 (first entry)
 XX
 DE Human DNA ligase III cDNA.
 XX
 KW DNA ligase III; tumour; cancer; immunosuppression; diagnosis;
 KW gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH FT 334..3102
 FT CDS /*lag- a
 FT 1810..1812
 FT /*lag- b
 FT /*transl_except- (1810..1812, aa:His)
 FT /*note- "CAG encodes His, as shown in Fig 1"
 FT 1813..1815
 FT /*lag- c
 FT /*transl_except- (1813..1815, aa:Lys)
 FT /*note- "AAG encodes Lys, as shown in Fig 1"
 PN W09630524-A1.
 XX
 PD 03-OCT-1996.
 XX
 PE 31-MAR-1995; 95WO-US03939.
 XX
 PR 31-MAR-1995; 95WO-US03939.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Haseltine WA, Wei Y, Yu G;
 XX

DR WPI: 1996-455369/45.
DR P-PSDB: AAM05295.
XX
PT New isolated human DNA Ligase III polynucleotide(s) - used to
PT develop prods. for diagnosis or treatment of disorders such as
XX tumours and severe immunosuppression
XX
PS Claim 6; Page 46; 78pp; English.
XX
CC A CDNA clone (AAT39622) codes for human DNA ligase III (AAM05295), an
CC enzyme that repairs single strand breaks in DNA. It was isolated
CC from a CDNA library derived from human testis, and may also be obtd.
CC from prostate, heart and thymus. The CDNA clone can be used to
CC detect mutant DNA ligase genes, and to produce recombinant DNA
CC ligase III in transformed host (e.g. bacterial, COS, insect) cells.
CC It can also be used in gene therapy to treat disorders associated
CC with a defect in DNA ligase III; antisense sequences can be used to
CC destroy undersized (e.g. cancer) cells.
XX
SQ Sequence 3417 BP; 899 A; 854 C; 912 G; 752 T; 0 other;

Query Match 72.4%; Score 15.2; DB 17; Length 3417;
Best Local Similarity 85.0%; Pred. No. 1.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 cataatggcctcgtcacac 21
DB 2275 CATCATGGCCTCTGCACAC 2256
||| ||||| ||||| |||||

RESULT 13
AAV82491/C
ID AAV82491 standard; CDNA: 3417 BP.
XX
AC AAV82491;
XX
DT 19-MAR-1999 (first entry)
XX
DE Human DNA ligase III encoding cDNA.
XX
KW Human; DNA ligase III; abnormal cellular proliferation; leukaemia;
KW cancer; tumour; immunosuppression; stunted growth; lymphoma;
KW cellular hypersensitivity; DNA-damaging agent; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 334..3102
FT /*tag= a
XX
PN US5858705-A.
XX
PD 12-JAN-1999.
XX
PF 05-JUN-1995; 95US-0464402.
XX
PR 05-JUN-1995; 95US-0464402.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Haseltine WA, Wei Y, Yu G;
XX
DR WPI: 1999-119875/10.
DR P-PSDB: AAM89575.
XX
PT DNA that codes for the DNA ligase III polypeptide - useful for
PT producing recombinant polypeptides
XX
PS Claim 14; Fig 1; 39pp; English.
XX
CC The present sequence encodes human DNA ligase III. Human DNA ligase III
CC is located within bands 1/911.2-12. The human DNA ligase III gene can be

CC used to treat conditions related to insufficient human DNA ligase III
CC activity via gene therapy by inserting the DNA ligase III into a
CC patients cells either in vivo or ex vivo. The gene is expressed in
CC transduced cells and as a result the protein may be used therapeutically
CC to prevent disorders associated with defects in DNA e.g. abnormal
CC cellular proliferation such as leukaemia, cancer and tumour; to treat
CC severe immunosuppression; stunted growth and lymphoma, as well as
CC cellular hypersensitivity to DNA-damaging agents.
XX
SQ Sequence 3417 BP; 900 A; 852 C; 912 G; 753 T; 0 other;

Query Match 72.4%; Score 15.2; DB 20; Length 3417;
Best Local Similarity 85.0%; Pred. No. 1.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 cataatggcctcgtcacac 21
DB 2275 CATCATGGCCTCTGCACAC 2256
||| ||||| ||||| |||||

RESULT 14
AAx77721/C
ID AAx77721 standard; DNA: 9775 BP.
XX
AC AAx77721;
XX
DT 10-AUG-1999 (first entry)
XX
DE N. crassa his-3 cogl 1pl region DNA.
XX
KW CogL: his-3; haploid; fungal cell; recombinant hot-spot; diploid;
KW diversification; sequence diversity; heterokaryon; transfection;
KW discontinuous conversion tract; frequency; ss.
XX
OS Neurospora crassa.
XX
PN WO9927072-A1.
XX
PD 03-JUN-1999.
XX
PF 23-NOV-1998; 98WO-AU00971.
XX
PR 24-NOV-1997; 97US-0977171.
XX
PA (FLIN-) FLINDERS TECHNOLOGIES PTY LTD.
XX
PI Catchside DE;
XX
DR WPI: 1999-357826/30.
XX
PT Haploid fungal cells comprising heterologous DNA coupled to a
PT recombination hot spot
XX
PS Example 4; Fig 7; 103pp; English.
XX

This invention describes a novel haploid fungal cell comprising a
CC recombinant genome, comprising a heterologous DNA functionally coupled
CC to a recombinant hot-spot. The haploid fungal cell is capable of being
CC converted to a diploid fungal cell, where the heterologous DNA is adapted
CC and configured in the recombinant genome for recombination in the diploid
CC fungal cell. The methods and compositions of the invention are useful for
CC the diversification of heterologous DNA sequences in vivo. The process
CC generates new versions of the foreign DNA by recombining their
CC differences in new combinations. Errors in recombination generate
CC additional sequence diversity. The strains containing diversified
CC heterologous sequences are used directly for expression of the variant
CC gene or/are combined in pairwise or higher order combinations in
CC heterokaryons where a heteromeric protein, such as an immunoglobulin is
CC the product. The DNA variants can be derived from genes of different
CC species having homologous genes that produce proteins with equivalent
CC function but differing properties such as pH tolerance, thermostability
CC and substrate range, or for monoclonal antibodies derived from different

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2001, 23:28:18 : Search time 4309.39 Seconds
(without alignments)
46.065 Million cell updates/sec

Title: US-09-138-735-10
Perfect score: 21
Sequence: 1 acataatgcctcgttcacac 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

```
117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*
```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|-----|-----------|--------------------|
| 1 | 17.8 | 84.8 | 508 | 238 | AZ109064 | RPCI-23-4 |
| 2 | 16.8 | 80.0 | 156 | 133 | BF171313 | PCL2451.M |
| 3 | 16.8 | 80.0 | 501 | 22 | A1588710 | fb98e05.y |
| 4 | 16.8 | 80.0 | 540 | 219 | CNS000CN | AL082829 Arabidops |
| 5 | 16.8 | 80.0 | 548 | 232 | AQ728515 | HS_5481_A |
| 6 | 16.8 | 80.0 | 551 | 123 | AW933866 | RC3-BN003 |
| 7 | 16.8 | 80.0 | 574 | 162 | BE019806 | b559003.y |
| 8 | 16.8 | 80.0 | 618 | 222 | FR0008179 | 29189.F.rubripes |
| 9 | 16.8 | 80.0 | 674 | 228 | AQ390364 | CTBTB-EL |
| 10 | 16.8 | 80.0 | 683 | 104 | A1958169 | fc91f07.y |
| 11 | 16.8 | 80.0 | 852 | 220 | CNS021D5 | AL213026 Tetradon |
| 12 | 16.8 | 80.0 | 882 | 221 | CNS03GLR | AL243144 Tetradon |
| 13 | 16.4 | 78.1 | 665 | 120 | AW766833 | da71g12.x |
| 14 | 16.4 | 78.1 | 937 | 169 | BF793589 | 602254242 |
| 15 | 16.2 | 77.1 | 135 | 21 | A1537376 | tp07h11.x |
| 16 | 16.2 | 77.1 | 175 | 23 | A1708599 | as62a04.x |
| 17 | 16.2 | 77.1 | 182 | 19 | A1351780 | q109b06.x |
| 18 | 16.2 | 77.1 | 199 | 124 | BB053373 | BB053373 |
| 19 | 16.2 | 77.1 | 211 | 24 | A1749861 | ac33d04.x |
| 20 | 16.2 | 77.1 | 235 | 120 | AW793754 | MRI-TM000 |
| 21 | 16.2 | 77.1 | 253 | 17 | A1198558 | qf49h06.x |
| 22 | 16.2 | 77.1 | 292 | 8 | AA483447 | ne65g09.s |
| 23 | 16.2 | 77.1 | 298 | 24 | A1749187 | at40d03.x |
| 24 | 16.2 | 77.1 | 302 | 104 | A1968537 | wt38d11.x |
| 25 | 16.2 | 77.1 | 310 | 23 | A1708151 | as38c05.x |
| 26 | 16.2 | 77.1 | 318 | 103 | A1866989 | wn14b02.x |
| 27 | 16.2 | 77.1 | 318 | 128 | BB208297 | BB208297 |
| 28 | 16.2 | 77.1 | 383 | 110 | AV741908 | AV741908 |
| 29 | 16.2 | 77.1 | 402 | 239 | A2157187 | SP_0022_A |
| 30 | 16.2 | 77.1 | 441 | 148 | BF431250 | naa42h09. |
| 31 | 16.2 | 77.1 | 447 | 142 | BE968472 | 601649532 |
| 32 | 16.2 | 77.1 | 458 | 21 | A1521229 | to66e09.x |
| 33 | 16.2 | 77.1 | 458 | 223 | AQ039278 | CIT-HSP-2 |
| 34 | 16.2 | 77.1 | 531 | 163 | BE115460 | UT-R-BS1 |
| 35 | 16.2 | 77.1 | 554 | 22 | A1571873 | AL1571873 |
| 36 | 16.2 | 77.1 | 559 | 23 | A1680156 | tw65e05.x |
| 37 | 16.2 | 77.1 | 571 | 156 | C18439 | C18439 |
| 38 | 16.2 | 77.1 | 595 | 9 | AA574062 | nk11d08.s |
| 39 | 16.2 | 77.1 | 640 | 146 | BF215276 | 601846322 |
| 40 | 16.2 | 77.1 | 643 | 168 | BF694071 | 602082675 |
| 41 | 16.2 | 77.1 | 673 | 32 | AV713711 | AV713711 |
| 42 | 16.2 | 77.1 | 708 | 24 | A1721846 | fc26a10.y |
| 43 | 16.2 | 77.1 | 717 | 110 | AV762205 | AV762205 |
| 44 | 16.2 | 77.1 | 774 | 147 | BF338092 | 602035967 |
| 45 | 16.2 | 77.1 | 826 | 168 | BF690968 | 602246993 |

ALIGNMENTS

| | | | | | |
|------------|--|------------|---------|----------------------------|----------------|
| RESULT 1 | AZ109064 | 508 bp | DNA | GSS | 09-MAY-2000 |
| LOCUS | RPCI-23-471L14 | TV | RPCI-23 | Mus musculus genomic clone | RPCI-23-471L14 |
| DEFINITION | , DNA sequence. | | | | |
| ACCESSION | AZ109064 | | | | |
| KEYWORDS | AZ109064.1 | GI:7762122 | | | |
| SOURCE | house mouse. | | | | |
| ORGANISM | Mus musculus | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. | | | | |
| AUTHORS | Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Alinet,B., Levins,M., McGann,S., Isegaye,G., Geer,K., Kroll,M., de Jong,P. and Fraser,C.M. | | | | |
| TITLE | Mouse BAC End Sequences from Library RPCI-23 | | | | |

JOURNAL COMMENT

Unpublished (1999)
Other GSSs: RPCI-23-471L14.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhaoc@igf.org

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderframe.htm>) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 471 row: L column: 14
Seq primer: SP6
Class: BAC ends.

FEATURES

source

1..508

Location/Qualifiers

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-23-471L14"

/clone_lib="RPCI-23"

/sex="Female"

/lab_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site:1: EcoRI; Site:2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Site selected DNA was cloned into the pBAC3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT

120 a 91 c 102 g 195 t

Query Match 84.8%; Score 17.8; DB 238; Length 508;
Best Local Similarity 90.5%; Pred. No. 76;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 acataatggccttcacac 21
|||||
Db 215 ACATGATGCCCTTTCACAC 195

| | | | | | |
|------------|---|---------------|--------------|--------------|-------------|
| RESULT 2 | BF171313 | 156 bp | mRNA | EST | 23-MAR-2001 |
| LOCUS | PCL2451 | Myeloma (PCL) | CDNA Library | Homo sapiens | CDNA, mRNA |
| DEFINITION | sequence. | | | | |
| ACCESSION | BF171313 | | | | |
| KEYWORDS | BF171313.1 | GI:13437527 | | | |
| SOURCE | human. | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | |
| AUTHORS | 1 (bases 1 to 156) Claudio,J.O., Tang,H., Khan,E.M., Voralia,M., Li,Z., Cukerman,E., Francisco-Pabalan,O., Liew,C.C. and Stewart,A.K. | | | | |
| TITLE | The transcriptional phenotype of myeloma cells | | | | |
| JOURNAL | Unpublished (2000) | | | | |
| COMMENT | Contact: A. Keith Stewart, M.D. Oncology Research University Health Network 610 University Ave., 5-126, Toronto, Ontario, M5G 2M9, Canada Tel: (416) 946-4639 Fax: (416) 946-6546 Email: k.stewart@utoronto.ca PCR Primers FORWARD: 5'-GCCAGCTCGAATTAACCTTCACCTAAAGG-3' | | | | |

BACKWARD: 5'-CCAGTGATTTGTAATACGACCTCATTATAGGCC-3'
Seq primer: 5'-GAAATTAACCTCCTCAAAAGG-3'.
Location/Qualifiers

FEATURES

source

1. 156
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Myeloma (PCL) cDNA library"
/sex="male"
/tissue_type="Blood"
/cell_type="myeloma"
/dev_stage="Plasma cell leukemia"
/note="Vector: lambda Zap Express; Site_1: EcoRI; Site_2: XhoI; mRNA was purified from plasma cell leukemia patient's peripheral blood containing >95% myeloma. An oligo d(T)18 primer containing XhoI restriction site was used to prime first strand synthesis using M-MLV reverse transcriptase. To protect the cDNAs from XhoI digestion in subsequent cloning step, the nucleotide analogue 5-methyl-dCTP was added to the nucleotide mixture and a-32P dATP was added to monitor the quantity and quality of first strand synthesis. After second-strand synthesis and blunting of cDNA termini, EcoRI adapters were ligated, followed by kinase treatment and digestion with XhoI. The cDNAs were then size-fractionated using Sephacryl S-500 column and then ligated into EcoRI and XhoI digested Lambda Zap Express vector. The ligation product was packaged using Gigapack II packaging extract. The library had primary titre of approx. 1x10⁶. Clones from the primary library were randomly selected for single pass sequencing."

BASE COUNT 34 a 41 C 48 g 33 t
ORIGIN

Query Match 80.0%; Score 16.8; DB 145; Length 156;
Best Local Similarity 90.0%; Pred. No. 2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 cataatgacctcggtcacac 21
|||||
Db 148 CATATGCGCTCCTGCACAC 129

RESULT 3

AT588710/c 501 bp mRNA EST 21-APR-1999
LOCUS
DEFINITION
FB98605.y1 zebrafish washu mpimg EST Danio rerio cDNA 5' similar to
SW:EMT1_RAT P24942 EXCITATORY AMINO ACID TRANSPORTER 1 ; mRNA
sequence.

ACCESSION AT588710.1 GI:4597757

VERSION AT588710.1
KEYWORDS
SOURCE EST.
ORGANISM zebrafish.
Danio rerio

REFERENCE
AUTHORS
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Actinopterygii: Neopterygii: Teleostei: Euteleostei: Ostariophysi:
Cypriniformes: Cyprinidae: Rasbora: Danio.
1 (bases 1 to 501)

Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy,
S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood,
K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B.,
Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E.,
Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.
and Wilson, R.

Washu Zebrafish EST Project 1998
Unpublished (1998)

CONTACT: Stephen L. Johnson

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: zbrfish@wustl.edu

CDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:

FEATURES

source

Matthew Clark. DNA Sequencing by: Washington University Genome
Sequencing Center Clone distribution: Genome Systems, St. Louis,
Missouri (web address: www.genomesystems.com) (email contact:
info@genomesystems.com) and Research Genetics, Huntsville, Alabama
(web address: www.resgen.com) (email contact: info@resgen.com) and
Ressourcen/Zentrum/Primardatenbank, Berlin, Germany (web address:
www.rzpd.de)
Seq primer: T3 ET from Amersham
High quality sequence stop: 477.
Location/Qualifiers

1. 501
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone_lib="Zebrafish Washu MPIMG EST"
/sex="mixed"
/tissue_type="26 somite embryos, adult livers, shield
stage embryos"
/lab_host="X11-Blue MP"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st
strand cDNA was primed with a Not I - oligo(dt)18 primer
[5'-GACATGTTCTAGATCGCGAGCGCGCCCTTTTCTTTT3'];
double-stranded cDNA was ligated to Sal I adaptors (BRL),
digested with Not I and cloned into the Not I and Sal I
sites of the pSPORT1 vector (BRL). Library was constructed
by Matthew Clark (Lehrach lab, ICRF, London and Max Planck
Institut fuer Molekulare Genetik, Berlin). cDNAs for EST
analysis were selected following oligonucleotide
hybridization fingerprinting of arrayed clones from
zebrafish late somitogenesis (26 ss), adult liver or
embryonic shield stage (5.6 h) libraries. Fingerprint
data were used to computationally cluster cDNAs, and a
single cDNA from each cluster was chosen for sequencing.
In some cases multiple members of the same cluster were
sequenced to assess clustering parameters or single clones
were sequenced additional times to assess quality
control."

BASE COUNT 134 a 122 C 131 g 114 t
ORIGIN

Query Match 80.0%; Score 16.8; DB 22; Length 501;
Best Local Similarity 90.0%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 cataatgacctcggtcacac 21
|||||
Db 464 CATGATGCGCTCCTGCAGAC 445

RESULT 4

CNS0006N 540 bp DNA GSS 28-JUN-1999
LOCUS
DEFINITION
Arabidopsis thaliana genome survey sequence 77 end of BAC F5J23 of
IGF library from strain Columbia of Arabidopsis thaliana, genomic
survey sequence.

ACCESSION AL082829.1 GI:5283969

VERSION AL082829.1
KEYWORDS
SOURCE
ORGANISM

GSS.
thale cress.
Arabidopsis thaliana
Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:
Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eurosids II:
Brassicales: Brassicaceae: Arabidopsis.

REFERENCE
AUTHORS
1 (bases 1 to 540)

Salamounab, M., Choisme, N., Artiguenave, F., Brotlier, P., Wincker, P.,
Samson, D., Saurin, W., Weissenbach, J. and Quetier, F.

Unpublished

2 (bases 1 to 540)

Genoscope.

Direct Submission

Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

FEATURES
Source
Location/Qualifiers
1. .548
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone_lib="IGF"
/clone="F5J23"
/note="end : T7"
BASE COUNT 171 a 105 c 104 g 160 t
ORIGIN

Query Match 80.0%; Score 16.8; DB 219; Length 540;
Best Local Similarity 90.0%; Pred. No. 2.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 cataatgacctgtcacac 21
|||||
Db 321 CATATAGCGCTGTTACTC 340

RESULT 5
A0728515 548 bp DNA GSS 15-JUL-1999
LOCUS HS_5481_A2_F05_77A RPCI-11 Human Male BAC Library Homo sapiens
DEFINITION genomic clone Plate-1057 Col-10 Row-K, DNA sequence.
ACCESSION A0728515
VERSION A0728515.1 GI:5500067
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 548)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURN. Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (<http://www.htsc.washington.edu>)
<http://www.htsc.washington.edu>
Plate: 1057 row: K column: 10
Seq primer: T7
Class: BAC ends
High quality sequence stop: 548.
Location/Qualifiers
1. .548
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=1057 Col=10 Row=K"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="Male"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACe3.6 vector at EcoRI sites"
BASE COUNT 167 a 114 c 108 g 148 t 11 others
ORIGIN

Query Match 80.0%; Score 16.8; DB 232; Length 548;
Best Local Similarity 90.0%; Pred. No. 2.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 acataatgacctgtcacac 20
|||||
Db 425 ACATATAGCGCTGTTACACA 444

RESULT 6
AW93866 551 bp mRNA EST 05-JUN-2000
LOCUS RC3-BN0034-240400-017-a07 BN0034 Homo sapiens CDNA, mRNA sequence.
DEFINITION
ACCESSION AW93866
VERSION AW93866.1 GI:8254053
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 551)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURN. Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=6t2-RC3-BN0034-240>)
400-017-a07&t3=2000-04-24&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 15
High quality sequence stop: 551.
Location/Qualifiers
1. .551
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BN0034"
/dev_stage="Adult"
/note="Organ: breast_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 157 a 123 c 100 g 171 t
ORIGIN

Query Match 80.0%; Score 16.8; DB 123; Length 551;
Best Local Similarity 90.0%; Pred. No. 2.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 acataatgacctgtcacac 20
|||||
Db 258 ACATATAGCGCTGTTACACA 277

```

RESULT 7
LOCUS BE019806/c 574 bp mRNA EST 06-JUN-2000
DEFINITION bb59h03.y1 NIH.MGC.9 Homo sapiens cDNA clone IMAGE:3029813 5',
similar to SW:RS_MOUSE P97461 40S RIBOSOMAL PROTEIN S5. ;, mRNA
sequence.
ACCESSION BE019806
VERSION BE019806.1 GI:8279893
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 574)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@b5-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
Image.lnl.gov/Image/html/lresources.shtml
Seq primer: -40RP from Gibco
High quality sequence stop: 337.
location/Qualifiers
1. 574
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1ib="NIH-MGC.9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 133 a 105 c 165 g 110 t 1 others
ORIGIN

Query Match 80.0%; Score 16.8; DB 162; Length 574;
Best Local Similarity 90.0%; Pred. No. 2.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 cataatgacctgtcacac 21
|||||
Db 481 CATGATGCGCTGTTACAC 462

RESULT 8
LOCUS FR0008179/c 618 bp DNA GSS 02-MAR-1997
DEFINITION F.rubripes GSS sequence, clone 188G03AD3, genomic survey sequence.
ACCESSION 291989
VERSION 291989.1 GI:1869203
KEYWORDS GSS: genome survey sequence.
SOURCE Takifugu rubripes.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Takifugu.
1 (bases 1 to 618)
REFERENCE Elgar,G., Clark,M., Smith,S., Week,S., Warner,S., Umranta,Y.,
Williams,G. and Brenner,S.

```

```

TITLE Direct Submission
JOURNAL Submitted (18-FEB-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biolhelp@hgm.mrc.ac.uk
COMMENT Vector: m13mp18
V-type: phage
PRIMER: M13
DESCR: One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
source location/Qualifiers
1. 618
/organism="Takifugu rubripes"
/db_xref="taxon:31033"
/clone_1ib="cosmid 188G03"
/clone_1ib="188G03AD3"
BASE COUNT 161 a 154 c 145 g 152 t 6 others
ORIGIN

Query Match 80.0%; Score 16.8; DB 222; Length 618;
Best Local Similarity 90.0%; Pred. No. 2.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 cataatgacctgtcacac 21
|||||
Db 269 CCTAATGCGCTGTTACAC 250

RESULT 9
LOCUS AO390364/c 674 bp DNA GSS 06-MAR-1999
DEFINITION CITBI-E1-2545M12.TF CITBI-E1 Homo sapiens genomic clone 2545M12,
DNA sequence.
ACCESSION AO390364
VERSION AO390364.1 GI:4361387
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 674)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Ventur,J.C
TITLE Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: CITBI-E1-2545M12.TF
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES
source location/Qualifiers
1. 674
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1ib="2545M12"
/clone_1ib="CITBI-E1"
/sex="male"
/cell_type="sperm"
/note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"
BASE COUNT 200 a 138 c 162 g 174 t
ORIGIN

```


RESULT 12
LOCUS CNS03GLR 882 bp DNA GSS 17-MAY-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone
024H22 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL243144 GI:7964156
VERSION AL243144.1
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 882)
AUTHORS Roest-Crollius, H., Jallion, O., Dasilva, C., Fizames, C., Fisher, C.,
Bonneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
Weissenbach, J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 882)
AUTHORS Roest-Crollius, H., Jallion, O., Dasilva, C., Bonneau, L., Fisher, C.,
Bernot, A., Fizames, C., Wincker, P., Brothier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.
TITLE Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 882)
AUTHORS Direct Submission
TITLE Submitted (12-Apr-2000) to the EMBL/GenBank/DBJ databases
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
source
1..882
Location/Qualifiers
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="024H22"
/clone_1lb="G"
/note="Genoscope sequence ID : COB024DD11LPI-end : T7"
BASE COUNT 176 a 204 c 200 g 300 t 2 others
ORIGIN
Query Match 80.0%; Score 16.8; DB 221; Length 882;
Best Local Similarity 90.0%; Pred. No. 2.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2 cataatggcctcggtcacac 21
||||| |||||||
Db 270 CATCATGCTCTGCTTCACAC 289
RESULT 13
LOCUS AM766833 665 bp mRNA EST 16-FEB-2001
DEFINITION da71g12.x1 Harland strage 19-23 Xenopus laevis cDNA clone
IMAGE:3200422.3, similar to SW:CA24_MOUSE P08122 COLLAGEN ALPHA
2(IV) CHAIN PRECURSOR. [1] ;, mRNA sequence.
ACCESSION AM766833 GI:7698826
VERSION AM766833.1
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 665)
AUTHORS Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D.

TITLE
JOURNAL
COMMENT
Martlin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person
, B., Gibbons, M., Harvey, N., Rifter, E., Jackson, Y., McCann, R.,
Waterson, R. and Wilson, R.
Washu Xenopus EST project, 1999
Unpublished (1999)
Contact: Sandy Clifton, Ph.D.
Washu Xenopus EST project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estelawson.wustl.edu
Library constructed by R. Harland, PhD (University of California,
Berkeley)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/tresources.shtml
Seq primer: -400P from Gibco
High quality sequence stop: 459.
Location/Qualifiers
1..665
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="IMAGE:3200422"
/clone_1lb="Harland stage 19-23"
/tissue_type="neurula"
/dev_stage="stage 19-23"
/lab_host="DH10B (phage-resistant)"
/note="Vector: PCS107 (custom); Site_1: NotI; Site_2: SalI
; cDNA made by oligo-dT priming. Library constructed by
Dr. Francesca Mariani in the laboratory of R. Harland,
Ph.D. (University of California, Berkeley). References:
XBF-2 is a transcriptional repressor that converts
ectoderm into neural tissue. Mariani, FV, Harland, RM.,
development. 1998 Dec;125(24):5019-31. PMID: 9811586; UI:
99030283; Use of large-scale expression cloning screens in
the Xenopus laevis tadpole to identify gene function.
Gramer TC, Liu KJ, Mariani FV, Harland RM., Dev Biol.
2000 Dec 15;228(2):197-210. PMID: 11112324; UI: 20564075;
Note: This is a Xenopus Gene Collection (XGC) library."
BASE COUNT 145 a 219 c 123 g 178 t
ORIGIN
Query Match 78.1%; Score 16.4; DB 120; Length 665;
Best Local Similarity 94.4%; Pred. No. 4.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 cataatggcctcggtcac 19
||||| |||||||
Db 104 CATATATGCTCTTTCAC 121
RESULT 14
LOCUS BF793589 937 bp mRNA EST 12-JAN-2001
DEFINITION 602254242F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4346435 5',
mRNA sequence.
ACCESSION BF793589
VERSION BF793589.1 GI:12098643
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 937)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM9968 row: c column: 12
High quality sequence stop: 667.
Location/Qualifiers

FEATURES

source

1. .937
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4346435"
/clone_id="NIH_MGC_84"
/tissue_type="adrenal cortex carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: adrenal gland; Vector: PCMV-SPORT6; Site_1:
NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT
primed. Average insert size 1.229 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 262 a 168 c 182 g 325 t
ORIGIN

Query Match 78.1%; Score 16.4; DB 169; Length 937;
Best Local Similarity 94.4%; Pred. No. 4.6e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 taatgacctgctcacac 21
|||
Db 846 TATGGCCTCCTTCACAC 863

RESULT 15
AI537376/c 135 bp mRNA EST 13-MAY-1999
LOCUS tp07h11.x1 NCI-CGAP-Gas4 Homo sapiens CDNA IMAGE:2187141.3'
DEFINITION similar to TR:021697 021697 NADH DEHYDROGENASE SUBUNIT 4. ;, mRNA
sequence.
ACCESSION AI537376 GI:4451511
VERSION AI537376
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 135)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/bdrrp/image/image.html
Insert Length: 449 Std Error: 0.00
Seq primer: -40UP from GIBCO
High quality sequence stop: 50
POLYA-No.

FEATURES

source

Location/Qualifiers
1. .135
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2187141"
/clone_id="NCI-CGAP_Gas4"
/tissue_type="poorly differentiated adenocarcinoma with

signet ring cell features"
/lab_host="DH10B"
/note="Organ: stomach; Vector: PCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.69 kb. Life Technologies catalog #:
11549-011"
BASE COUNT 33 a 17 c 50 g 35 t
ORIGIN

Query Match 77.1%; Score 16.2; DB 21; Length 135;
Best Local Similarity 85.7%; Pred. No. 4e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 acataatgacctgctcacac 21
|||||
Db 104 ACATAAGCCCTCAATTCACAC 84

Search completed: September 20, 2001, 23:28:33
Job time: 25573 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 02:16:39 : Search time 164.23 Seconds
(without alignments)
24.207 Million cell updates/sec

Title: US-09-138-735-10

Perfect score: 21

Sequence: 1 acataatgcctcgttcacac 21

Scoring table: IDENTITY_NNC

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn1_7/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgn1_7/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn1_7/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn1_7/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn1_7/ptodata/1/ina/PCtUS.COMB.seq:*
- 6: /cgn1_7/ptodata/1/ina/backfillseq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 21 | 100.0 | 21 | 1 | US-08-480-917-10 |
| 2 | 21 | 100.0 | 3402 | 1 | US-08-480-917-1 |
| 3 | 15.8 | 75.2 | 2341 | 1 | US-08-441-430-2 |
| 4 | 15.8 | 75.2 | 3147 | 1 | US-08-441-430-3 |
| 5 | 15.8 | 75.2 | 4488 | 1 | US-08-441-430-1 |
| 6 | 15.2 | 72.4 | 3417 | 2 | US-08-464-402-1 |
| 7 | 15.2 | 72.4 | 9775 | 4 | US-08-977-171-1 |
| 8 | 15.2 | 72.4 | 9934 | 4 | US-08-977-171-2 |
| 9 | 14.8 | 70.5 | 2126 | 2 | US-08-789-354-1 |
| 10 | 14.8 | 70.5 | 2126 | 3 | US-09-110-937-1 |
| 11 | 14.8 | 70.5 | 2126 | 3 | US-09-058-725B-1 |
| 12 | 14.8 | 70.5 | 2126 | 3 | US-09-232-857-1 |
| 13 | 14.8 | 70.5 | 4465 | 2 | US-08-620-605D-1 |
| 14 | 14.8 | 70.5 | 4547 | 2 | US-09-005-232A-1 |
| 15 | 14.6 | 69.5 | 338 | 4 | US-08-991-789A-82 |
| 16 | 14.6 | 69.5 | 501 | 2 | US-08-454-557C-34 |
| 17 | 14.6 | 69.5 | 501 | 2 | US-08-340-426D-34 |
| 18 | 14.6 | 69.5 | 501 | 2 | US-08-450-673C-34 |
| 19 | 14.6 | 69.5 | 501 | 5 | PCT-US95-17111A-34 |
| 20 | 14.6 | 69.5 | 594 | 2 | US-08-454-557C-29 |
| 21 | 14.6 | 69.5 | 594 | 2 | US-08-340-426D-29 |
| 22 | 14.6 | 69.5 | 594 | 2 | US-08-450-673C-29 |
| 23 | 14.6 | 69.5 | 594 | 5 | PCT-US95-17111A-29 |
| 24 | 14.6 | 69.5 | 16569 | 4 | US-09-097-889-2 |
| 25 | 14.2 | 67.6 | 270 | 2 | US-08-921-382-16 |
| 26 | 14.2 | 67.6 | 1158 | 4 | US-09-198-092-1 |
| 27 | 14.2 | 67.6 | 3240 | 4 | US-09-262-773-7 |

| | | | | | | |
|----|------|------|-------|---|--------------------|-------------------|
| 28 | 14.2 | 67.6 | 3244 | 4 | US-09-262-773-3 | Sequence 3, Appl |
| 29 | 14.2 | 67.6 | 3264 | 4 | US-09-262-773-5 | Sequence 5, Appl |
| 30 | 14.2 | 67.6 | 3268 | 4 | US-09-262-773-1 | Sequence 1, Appl |
| 31 | 14.2 | 67.6 | 4652 | 2 | US-07-861-800-1 | Sequence 1, Appl |
| 32 | 14.2 | 67.6 | 20137 | 4 | US-09-262-773-206 | Sequence 206, App |
| 33 | 14.2 | 67.6 | 20138 | 4 | US-09-262-773-9 | Sequence 9, Appl |
| 34 | 14.2 | 67.6 | 23071 | 4 | US-09-262-773-210 | Sequence 210, App |
| 35 | 13.8 | 65.7 | 491 | 1 | US-08-133-711-38 | Sequence 38, Appl |
| 36 | 13.8 | 65.7 | 2047 | 4 | US-09-345-468-1 | Sequence 1, Appl |
| 37 | 13.8 | 65.7 | 2520 | 1 | US-08-087-016-1 | Sequence 1, Appl |
| 38 | 13.6 | 64.8 | 384 | 3 | US-09-284-782-33 | Sequence 33, Appl |
| 39 | 13.6 | 64.8 | 554 | 2 | US-08-454-557C-31 | Sequence 31, Appl |
| 40 | 13.6 | 64.8 | 554 | 2 | US-08-340-426D-31 | Sequence 31, Appl |
| 41 | 13.6 | 64.8 | 554 | 2 | US-08-450-673C-31 | Sequence 31, Appl |
| 42 | 13.6 | 64.8 | 554 | 5 | PCT-US95-17111A-31 | Sequence 31, Appl |
| 43 | 13.6 | 64.8 | 569 | 2 | US-08-485-778-5 | Sequence 5, Appl |
| 44 | 13.6 | 64.8 | 1296 | 5 | PCT-US91-00909-5 | Sequence 5, Appl |
| 45 | 13.6 | 64.8 | 1425 | 3 | US-09-009-494-5 | Sequence 5, Appl |

ALIGNMENTS

RESULT 1
US-08-480-917-10
Sequence 10, Application US/08480917
Patent No. 5620864
GENERAL INFORMATION:
APPLICANT: PARANHOS-BACCALA, Glaucia
APPLICANT: LESENECHAL, Mylene
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: NEW TRYPAANOSOMA CRUZI ANTIGEN, AND GENE
TITLE OF INVENTION: ENCODING THE LATTER, THEIR APPLICATION TO THE DETECTION OF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Oliff & Berridge
STREET: 700 South Washington Street, Suite 300
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,917
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: MPB 36400
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-480-917-10
Query Match 100.0% Score 21; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.02; Mismatches 0; Gaps 0;
Matches 21; Conservative 0; Indels 0;

Db 1 ACATATGCGCTCGTCACAC 21

RESULT 2

US-08-480-917-1/C
Sequence 1, Application US/08480917
Patent No. 5820864
GENERAL INFORMATION:
APPLICANT: PARANHOS-BACALA, Glaucia
APPLICANT: LESENECHAL, Mylene
APPLICANT: JOLIVET, Michel
TITLE OF INVENTION: NEW TRYANOSOMA CRUZI ANTIGEN, AND GENE
TITLE OF INVENTION: ENCODING THE LATTER; THEIR APPLICATION TO THE DETECTION OF
TITLE OF INVENTION: CHAGAS DISEASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Oliff & Berridge
STREET: 700 South Washington Street, Suite 300
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,917
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Berridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: MPB 36400
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3402 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-480-917-1

Query Match 100.0%; Score 21; DB 1; Length 3402;

Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 acataatgacctcgttcacac 21
Db 2207 ACATATGCGCTCGTCACAC 2187

RESULT 3

US-08-441-430-2/C
Sequence 2, Application US/08441430
Patent No. 5681942
GENERAL INFORMATION:
APPLICANT: Buchwald, Manuel
APPLICANT: Strathdee, Craig A.
APPLICANT: Wevick, Rachel
APPLICANT: Mathew, Christopher George Porter
TITLE OF INVENTION: Fanconi Anemia Type C Gene
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Richard J. Polley, Esq.
ADDRESSEE: Klarquist, Sparkman, Campbell, Leigh &
ADDRESSEE: Winston, LLP

STREET: 121 S.W. Salmon, Suite 1600
CITY: Portland
STATE: Oregon
COUNTRY: U.S.A.
ZIP: 97204

COMPUTER READABLE FORM:
MEDIUM TYPE: Disk, 3+1/2-inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: WordPerfect 5.1/ASCII Text File
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,430
FILING DATE: May 15, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 07/876,285
FILING DATE: April 29, 1992
APPLICATION NUMBER: U.S. 07/918,313
FILING DATE: July 21, 1992
APPLICATION NUMBER: U.S. 08/003,963
FILING DATE: January 15, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Richard J. Polley, Esq.
REGISTRATION NUMBER: 28,107
REFERENCE/DOCKET NUMBER: 3812-42824
TELECOMMUNICATION INFORMATION:
TELEPHONE: (503) 226-7391
TELEFAX: (503) 228-9446
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2341 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double stranded
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: Human CDNA
POSITION IN GENOME: (of corresponding genomic
POSITION IN GENOME: gene)
CHROMOSOME/SEGMENT: 9q
MAP POSITION: 22.3
UNITS:
US-08-441-430-2

Query Match 75.2%; Score 15.8; DB 1; Length 2341;

Best Local Similarity 89.5%; Pred. No. 21;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 cataatgacctcgttcaca 20
Db 939 CAAATGCGCTCGTTTACA 921

RESULT 4

US-08-441-430-3/C
Sequence 3, Application US/08441430
Patent No. 5681942
GENERAL INFORMATION:
APPLICANT: Buchwald, Manuel
APPLICANT: Strathdee, Craig A.
APPLICANT: Wevick, Rachel
APPLICANT: Mathew, Christopher George Porter
TITLE OF INVENTION: Fanconi Anemia Type C Gene
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Richard J. Polley, Esq.
ADDRESSEE: Klarquist, Sparkman, Campbell, Leigh &
ADDRESSEE: Winston, LLP

MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,402
FILING DATE: June 5, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/03939
FILING DATE: 31 MAR 95
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-388
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3417 BASE PAIRS
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
US-08-464-402-1

Query Match 72.4%; Score 15.2; DB 2; Length 3417;
Best Local Similarity 85.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 cataatgacctcgttcacac 21
||| ||||| |||||
Db 2275 CATCATGCGCTCCTGCACAC 2256

RESULT 7
US-08-977-171-1/c
; Sequence 1, Application US/08977171
; Patent No. 6232112
; GENERAL INFORMATION:
; APPLICANT: CATCHESIDE, DAVID E.
; TITLE OF INVENTION: REAGENTS AND METHODS FOR DIVERSIFICATION
; TITLE OF INVENTION: OF DNA
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 6232112west Center, 90 South 7th Street
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/977,171
; FILING DATE: 24-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: SKOOG, MARK T
; REGISTRATION NUMBER: 40,178
; REFERENCE/DOCKET NUMBER: 10552.13US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9775 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-977-171-1

Query Match 72.4%; Score 15.2; DB 4; Length 9775;
Best Local Similarity 85.0%; Pred. No. 55;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 cataatgacctcgttcacac 21
||| ||||| |||||
Db 1397 CCGAGGCGCTCCTGCACAC 1378

RESULT 8
US-08-977-171-2/c
; Sequence 2, Application US/08977171
; Patent No. 6232112
; GENERAL INFORMATION:
; APPLICANT: CATCHESIDE, DAVID E.
; TITLE OF INVENTION: REAGENTS AND METHODS FOR DIVERSIFICATION
; TITLE OF INVENTION: OF DNA
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 6232112west Center, 90 South 7th Street
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/977,171
; FILING DATE: 24-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: SKOOG, MARK T
; REGISTRATION NUMBER: 40,178
; REFERENCE/DOCKET NUMBER: 10552.13US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9934 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-977-171-2

Query Match 72.4%; Score 15.2; DB 4; Length 9934;
Best Local Similarity 85.0%; Pred. No. 55;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 cataatgacctcgttcacac 21
||| ||||| |||||
Db 1501 CCGAGGCGCTCCTGCACAC 1482

```

RESULT 9
US-08-789-354-1/c
: Sequence 1, Application US/08789354
: Patent No. 5851798
: GENERAL INFORMATION:
: APPLICANT: Shabon, Usman
: APPLICANT: Bergsma, Derk
: TITLE OF INVENTION: Cloning of Human GPR14 Re
: TITLE OF INVENTION: Ceptor
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SmithKline Beecham Corporation
: STREET: 709 Swedeland Road
: CITY: King of Prussia
: STATE: PA
: COUNTRY: USA
: ZIP: 19406
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/789,354
: FILING DATE: 27-JAN-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Han, William T
: REGISTRATION NUMBER: 34,344
: REFERENCE/DOCKET NUMBER: P50610
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-270-5219
: TELEFAX: 610-270-4026
: TELEX:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2126 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
:
US-08-789-354-1

Query Match 70.5%; Score 14.8; DB 2; Length 2126;
Best Local Similarity 88.9%; Pred. No. 69;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 taatgacctgattcaac 21
1 | | | | | | | | | | | | | |
bb 558 TGATGGCCTCGTCCAC 541

RESULT 10
US-09-110-937-1/c
: Sequence 1, Application US/09110937A
: Patent No. 6005074
: GENERAL INFORMATION:
: APPLICANT: SHABON, USMAN
: APPLICANT: BERGSMÄ, DERK
: TITLE OF INVENTION: CLONING OF HUMAN GPR14 RECEPTOR
: FILE REFERENCE: P50610-1
: CURRENT APPLICATION NUMBER: US/09/110,937A
: CURRENT FILING DATE: 1998-07-06
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 2126
: TYPE: DNA

```

```

US-09-110-937-1
ORGANISM: 'HOMO SAPIENS'
Query Match          70.5%; Score 14.8; DB 3; Length 2126;
Best Local Similarity 88.9%; Pred. No. 69;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      4    taatgacctgttcacac 21
         | ||||| ||||| |||
Db       558  TGATGGCCTCGTCCAC 541

RESULT 11
US-09-058-725B-1/C
Sequence 1, Application US/09058725B
Patent No. 613420
GENERAL INFORMATION:
APPLICANT: Ames, Robert
APPLICANT: Sarau, Henry
APPLICANT: Foley, James
APPLICANT: Chamber, Jon
TITLE OF INVENTION: A Method of Finding Angonist
TITLE OF INVENTION: and Antagonist To Human and Rat GPR 14
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: <19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/058.725B
FILING DATE: April 10, 1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/789,354
FILING DATE: 27-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T.
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: GP50005-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-09-058-725B-1

Query Match          70.5%; Score 14.8; DB 3; Length 2126;
Best Local Similarity 88.9%; Pred. No. 69;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      4    taatgacctgttcacac 21
         | ||||| ||||| |||
Db       558  TGATGGCCTCGTCCAC 541

RESULT 12
US-09-232-857-1/C
```

```

1  GENERAL INFORMATION:
2  APPLICANT: SHIBATANI, TAKEJI
3  APPLICANT: AKATSUKA, HIROYUKI
4  APPLICANT: KAWAI, ERI
5  TITLE OF INVENTION: GENE PARTICIPATING IN THE MECHANISM OF
6  TITLE OF INVENTION: SECRETION OF ESTERASE
7  NUMBER OF SEQUENCES: 4
8  CORRESPONDENCE ADDRESS:
9  ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
10 STREET: PO BOX 747
11 CITY: FALLS CHURCH
12 STATE: VA
13 COUNTRY: USA
14 ZIP: 22040-0747
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Floppy disk
17 COMPUTER: IBM PC compatible
18 OPERATING SYSTEM: PC-DOS/MS-DOS
19 SOFTWARE: Patentln Release #1.0, Version #1.30
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/08/620,605D
22 FILING DATE: 22-MAR-1996
23 CLASSIFICATION: 435
24 ATTORNEY/AGENT INFORMATION:
25 NAME: SVENSSON, LEONARD R
26 REGISTRATION NUMBER: 30,330
27 REFERENCE/DOCKET NUMBER: 0020-3955
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: (703)-205-8000
30 TELEFAX: (703)-205-8050
31 INFORMATION FOR SEQ ID NO: 1:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH: 4465 base pairs
34 TYPE: nucleic acid
35 STRANDEDNESS: double
36 TOPOLOGY: linear
37 MOLECULE TYPE: genomic DNA
38 ORIGINAL SOURCE:
39 STRAIN: Serratia marcescens Sr41
40 US-08-620-605D-1
41
42 Query Match 70.5%; Score 14.8; DB 2; Length 4465;
43 Best Local Similarity 88.9%; Pred. No. 78;
44 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
45
46 QY 1 acataatgctcgttca 18
47 ||||| ||||| |||||
48 Db 1534 ACATGATGCGCTGCTTCA 1517
49
50 RESULT 14
51 US-09-005-232A-1/c
52 Sequence 1, Application US/09005232A
53 Patent No. 5922568
54 GENERAL INFORMATION:
55 APPLICANT: SHIBATANI, TAKEJI
56 APPLICANT: AKATSUKA, HIROYUKI
57 APPLICANT: KAWAI, ERI
58 TITLE OF INVENTION: GENE PARTICIPATING IN THE MECHANISM OF
59 TITLE OF INVENTION: SECRETION OF ESTERASE
60 NUMBER OF SEQUENCES: 4
61 CORRESPONDENCE ADDRESS:
62 ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
63 STREET: PO BOX 747
64 CITY: FALLS CHURCH
65 STATE: VA
66 COUNTRY: USA
67 ZIP: 22040-0747
68 COMPUTER READABLE FORM:
69 MEDIUM TYPE: Floppy disk
70 COMPUTER: IBM PC compatible
71 OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/005,232A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)-205-8000
TELEFAX: (703)-205-8050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4547 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
STRAIN: Serratia marcescens Str41
US-09-005-232A-1

Query Match 70.5%; Score 14.8; DB 2; Length 4547;
Best Local Similarity 88.9%; Pred. No. 79;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 acataatggcctcggttca 18
||| ||||| |||||
Db 1534 ACATGATGGCCTCGGTCA 1517

RESULT 15
US-08-991-789A-82
Sequence 82, Application US/08991789A
Patent No. 6225054
GENERAL INFORMATION:
APPLICANT: Fridakis, Tony N.
Smith, John M.
Reed, Steven G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF BREAST CANCER
NUMBER OF SEQUENCES: 292
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed IP Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,789A
FILING DATE: 11-Dec-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E. R.
REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 210121.419C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 338 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 82:
US-08-991-789A-82

Query Match 69.5%; Score 14.6; DB 4; Length 338;
Best Local Similarity 81.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 1 acataatggcctcggttcaac 21
||| ||||| ||||| |||||
Db 234 ACATAAACCTCATTCACAC 254

Search completed: September 21, 2001, 02:16:42
Job time: 28235 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 02:19:54 : Search time 6407.28 Seconds
(without alignments)
48.431 Million cell updates/sec

Title: US-09-138-735-10
Perfect score: 21
Sequence: 1 acataatgcctcgttcacac 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1415048 seqs, 7388405095 residues

Total number of hits satisfying chosen parameters: 28310096

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

```
1: /cgnl_7/ptodata/1/pna/PCRUS.COMB.seq:*
2: /cgnl_7/ptodata/1/pna/US06.COMB.seq:*
3: /cgnl_7/ptodata/1/pna/US07.COMB.seq:*
4: /cgnl_7/ptodata/1/pna/US08.COMB.seq:*
5: /cgnl_7/ptodata/1/pna/US081.COMB.seq:*
6: /cgnl_7/ptodata/1/pna/US082.COMB.seq:*
7: /cgnl_7/ptodata/1/pna/US083.COMB.seq:*
8: /cgnl_7/ptodata/1/pna/US084.COMB.seq:*
9: /cgnl_7/ptodata/1/pna/US085.COMB.seq:*
10: /cgnl_7/ptodata/1/pna/US086.COMB.seq:*
11: /cgnl_7/ptodata/1/pna/US087.COMB.seq:*
12: /cgnl_7/ptodata/1/pna/US088.COMB.seq:*
13: /cgnl_7/ptodata/1/pna/US089.COMB.seq:*
14: /cgnl_7/ptodata/1/pna/US090.COMB.seq:*
15: /cgnl_7/ptodata/1/pna/US091.COMB.seq:*
16: /cgnl_7/ptodata/1/pna/US092.COMB.seq:*
17: /cgnl_7/ptodata/1/pna/US093.COMB.seq:*
18: /cgnl_7/ptodata/1/pna/US094.COMB.seq:*
19: /cgnl_7/ptodata/1/pna/US095A.COMB.seq:*
20: /cgnl_7/ptodata/1/pna/US095B.COMB.seq:*
21: /cgnl_7/ptodata/1/pna/US095C.COMB.seq:*
22: /cgnl_7/ptodata/1/pna/US095D.COMB.seq:*
23: /cgnl_7/ptodata/1/pna/US096A.COMB.seq:*
24: /cgnl_7/ptodata/1/pna/US096B.COMB.seq:*
25: /cgnl_7/ptodata/1/pna/US096C.COMB.seq:*
26: /cgnl_7/ptodata/1/pna/US096D.COMB.seq:*
27: /cgnl_7/ptodata/1/pna/US096E.COMB.seq:*
28: /cgnl_7/ptodata/1/pna/US097A.COMB.seq:*
29: /cgnl_7/ptodata/1/pna/US097B.COMB.seq:*
30: /cgnl_7/ptodata/1/pna/US097C.COMB.seq:*
31: /cgnl_7/ptodata/1/pna/US098.COMB.seq:*
32: /cgnl_7/ptodata/1/pna/US098A.COMB.seq:*
33: /cgnl_7/ptodata/1/pna/US098B.COMB.seq:*
34: /cgnl_7/ptodata/1/pna/US098C.COMB.seq:*
35: /cgnl_7/ptodata/1/pna/US098D.COMB.seq:*
36: /cgnl_7/ptodata/1/pna/US098E.COMB.seq:*
37: /cgnl_7/ptodata/1/pna/US098F.COMB.seq:*
38: /cgnl_7/ptodata/1/pna/US098G.COMB.seq:*
39: /cgnl_7/ptodata/1/pna/US098H.COMB.seq:*
40: /cgnl_7/ptodata/1/pna/US098I.COMB.seq:*
41: /cgnl_7/ptodata/1/pna/US098J.COMB.seq:*
42: /cgnl_7/ptodata/1/pna/US098K.COMB.seq:*
43: /cgnl_7/ptodata/1/pna/US098L.COMB.seq:*
```

```
44: /cgnl_7/ptodata/1/pna/US6012.COMB.seq:*
45: /cgnl_7/ptodata/1/pna/US6013.COMB.seq:*
46: /cgnl_7/ptodata/1/pna/US6014.COMB.seq:*
47: /cgnl_7/ptodata/1/pna/US6015.COMB.seq:*
48: /cgnl_7/ptodata/1/pna/US6016.COMB.seq:*
49: /cgnl_7/ptodata/1/pna/US6017.COMB.seq:*
50: /cgnl_7/ptodata/1/pna/US6018.COMB.seq:*
51: /cgnl_7/ptodata/1/pna/US6019.COMB.seq:*
52: /cgnl_7/ptodata/1/pna/US6020.COMB.seq:*
53: /cgnl_7/ptodata/1/pna/US6021.COMB.seq:*
54: /cgnl_7/ptodata/1/pna/US6022.COMB.seq:*
55: /cgnl_7/ptodata/1/pna/US6023.COMB.seq:*
56: /cgnl_7/ptodata/1/pna/US6024.COMB.seq:*
57: /cgnl_7/ptodata/1/pna/US6025.COMB.seq:*
58: /cgnl_7/ptodata/1/pna/US6026.COMB.seq:*
59: /cgnl_7/ptodata/1/pna/US6027.COMB.seq:*
60: /cgnl_7/ptodata/1/pna/US6028.COMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 21 | 100.0 | 21 | 13 | US-08-988-242-6 |
| 2 | 21 | 100.0 | 21 | 13 | US-09-138-736-10 |
| 3 | 21 | 100.0 | 3402 | 15 | US-08-988-242-1 |
| 4 | 21 | 100.0 | 3402 | 15 | US-09-138-736-1 |
| 5 | 17.8 | 84.8 | 654 | 48 | US-60-167-217-21923 |
| 6 | 17.8 | 84.8 | 654 | 51 | US-60-191-637-21772 |
| 7 | 17.8 | 84.8 | 654 | 51 | US-60-191-637-21772 |
| 8 | 17.8 | 84.8 | 654 | 51 | US-60-191-637-21772 |
| 9 | 17.8 | 84.8 | 849 | 48 | US-60-173-464-17842 |
| 10 | 17.8 | 84.8 | 3382 | 48 | US-60-167-217-21922 |
| 11 | 17.8 | 84.8 | 3382 | 51 | US-60-191-637-21771 |
| 12 | 17.8 | 84.8 | 3382 | 51 | US-60-191-637-21771 |
| 13 | 17.8 | 84.8 | 3577 | 49 | US-60-173-464-17841 |
| 14 | 17.8 | 84.8 | 59140 | 19 | US-09-528-2374-1266 |
| 15 | 17.4 | 82.9 | 183411 | 53 | US-60-212-664-166 |
| 16 | 16.8 | 80.0 | 41001 | 56 | US-60-248-505-68 |
| 17 | 16.8 | 80.0 | 90824 | 20 | US-09-534-859-570 |
| 18 | 16.4 | 78.1 | 253 | 25 | US-09-654-617-160878 |
| 19 | 16.4 | 78.1 | 253 | 27 | US-09-684-016-160878 |
| 20 | 16.4 | 78.1 | 405 | 23 | US-09-605-699-20540 |
| 21 | 16.4 | 78.1 | 479 | 23 | US-09-605-699-20582 |
| 22 | 16.4 | 78.1 | 144838 | 53 | US-60-212-664-236 |
| 23 | 16.2 | 77.1 | 108 | 11 | US-08-798-074-1985 |
| 24 | 16.2 | 77.1 | 110 | 1 | PCT-US01-01334-1199 |
| 25 | 16.2 | 77.1 | 110 | 11 | US-08-798-074-1726 |
| 26 | 16.2 | 77.1 | 192 | 14 | US-09-628-858-604 |
| 27 | 16.2 | 77.1 | 265 | 52 | US-60-208-063-2221 |
| 28 | 16.2 | 77.1 | 267 | 51 | US-60-196-868-10176 |
| 29 | 16.2 | 77.1 | 392 | 25 | US-09-654-617-407207 |
| 30 | 16.2 | 77.1 | 392 | 27 | US-09-684-016-407207 |
| 31 | 16.2 | 77.1 | 478 | 19 | US-09-527-429-786 |
| 32 | 16.2 | 77.1 | 718 | 15 | US-09-644-871-6164 |
| 33 | 16.2 | 77.1 | 995 | 50 | US-60-182-316-937 |
| 34 | 16.2 | 77.1 | 2315 | 51 | US-60-173-464-9302 |
| 35 | 16.2 | 77.1 | 2315 | 51 | US-60-191-637-12205 |
| 36 | 16.2 | 77.1 | 2315 | 51 | US-60-191-637-12205 |
| 37 | 16.2 | 77.1 | 2318 | 48 | US-60-167-217-12199 |
| 38 | 16.2 | 77.1 | 4731 | 51 | US-60-173-464-9901 |
| 39 | 16.2 | 77.1 | 4731 | 51 | US-60-191-637-12204 |
| 40 | 16.2 | 77.1 | 4731 | 51 | US-60-191-637-12204 |
| 41 | 16.2 | 77.1 | 4734 | 48 | US-60-167-217-12198 |
| 42 | 16.2 | 77.1 | 8886 | 53 | US-09-620-392-41892 |
| 43 | 16.2 | 77.1 | 17116 | 24 | US-60-215-161-2406 |
| 44 | 16.2 | 77.1 | 17116 | 53 | US-60-215-161-2407 |
| 45 | 16.2 | 77.1 | 17116 | 53 | US-60-215-161-2408 |

ALIGNMENTS

```
RESULT 1
US-08-988-242-6
; Sequence 6, Application US/08988242
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: LESENECHAL, MYLENE
; APPLICANT: JOLIVET, MICHEL
; APPLICANT: MANDRAND, BERNARD
; TITLE OF INVENTION: NEW TRYPANOSOMA CRUZI ANTIGEN, GENE
; TITLE OF INVENTION: ENCODING THEREFOR, AND METHODS OF DETECTING AND TREATING
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIVIER & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/988,242
; FILING DATE: 10-DEC-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-988-242-6

Query Match          100.0%; Score 21; DB 13; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 acataatgacctgttcacac 21
    |||||||
Db 1 ACATAATGCGCTGTTCCACAC 21

RESULT 2
US-09-138-736-10
; Sequence 10, Application US/09138736
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: LESENECHAL, MYLENE
; APPLICANT: JOLIVET, MICHEL
; TITLE OF INVENTION: NEW TRYPANOSOMA CRUZI ANTIGEN, AND GENE
; TITLE OF INVENTION: ENCODING THE LATWER, THEIR APPLICATION TO THE DETECTION OF
; TITLE OF INVENTION: CHAGAS DISEASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Olivier & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
```

```
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/138,736
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/480,917
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-138-736-10
```

```
Query Match          100.0%; Score 21; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 acataatgacctgttcacac 21
    |||||||
Db 1 ACATAATGCGCTGTTCCACAC 21
```

```
RESULT 3
US-08-988-242-1/C
; Sequence 1, Application US/08988242
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: LESENECHAL, MYLENE
; APPLICANT: JOLIVET, MICHEL
; APPLICANT: MANDRAND, BERNARD
; TITLE OF INVENTION: NEW TRYPANOSOMA CRUZI ANTIGEN, GENE
; TITLE OF INVENTION: ENCODING THEREFOR, AND METHODS OF DETECTING AND TREATING
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIVIER & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/988,242
; FILING DATE: 10-DEC-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 36400A
```

TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3402 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-988-242-1

Query Match 100.0%; Score 21; DB 13; Length 3402;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 acataatgctcgttcacac 21
|||||
DB 2207 ACATAATGCGCTCGTTACAC 2187

RESULT 4
US-09-138-736-1/c
Sequence 1, Application US/09138736

GENERAL INFORMATION:
APPLICANT: PARANHOS-BACCALA, Glaucia
APPLICANT: LESENECHAL, Mylene
TITLE OF INVENTION: NEW TRYPANOSOMA CRUZI ANTIGEN, AND GENE
TITLE OF INVENTION: ENCODING THE LATTER, THEIR APPLICATION TO THE DETECTION OF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Oliff & Herridge
STREET: 700 South Washington Street, Suite 300
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/138,736
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/480,917
FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:
NAME: Herridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 36400
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3402 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-138-736-1

Query Match 100.0%; Score 21; DB 15; Length 3402;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 acataatgctcgttcacac 21
|||||
DB 2207 ACATAATGCGCTCGTTACAC 2187

RESULT 5
US-60-167-217-21923
Sequence 21923, Application US/60167217

GENERAL INFORMATION:
APPLICANT: Li, Peter W. D.
TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES
FILE REFERENCE: C1000152
CURRENT APPLICATION NUMBER: US/60/167,217
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 23195
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 21923
LENGTH: 654
TYPE: DNA
ORGANISM: Drosophila
US-60-167-217-21923

Query Match 84.8%; Score 17.8; DB 48; Length 654;
Best Local Similarity 90.5%; Pred. No. 73;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 acataatgctcgttcacac 21
|||||
DB 546 acataatgctcgttcacac 566

RESULT 6
US-60-191-637-21772
Sequence 21772, Application US/60191637

GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
FILE REFERENCE: C1000392
CURRENT APPLICATION NUMBER: US/60/191,637
CURRENT FILING DATE: 2000-03-23
NUMBER OF SEQ ID NOS: 42660
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 21772
LENGTH: 654
TYPE: DNA
ORGANISM: DROSOPHILA
US-60-191-637-21772

Query Match 84.8%; Score 17.8; DB 51; Length 654;
Best Local Similarity 90.5%; Pred. No. 73;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 acataatgctcgttcacac 21
|||||
DB 546 acataatgctcgttcacac 566

RESULT 7
US-60-191-681-17179
Sequence 17179, Application US/60191681

GENERAL INFORMATION:
APPLICANT: Li, Peter, W.D.
TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
FILE REFERENCE: C1000390
CURRENT APPLICATION NUMBER: US/60/191,681

;; CURRENT FILING DATE: 2000-03-23
;; NUMBER OF SEQ ID NOS: 30973
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 17179
;; LENGTH: 654
;; TYPE: DNA
;; ORGANISM: DROSOPHILA
US-60-191-681-17179

Query Match 84.8%; Score 17.8; DB 51; Length 654;
Best Local Similarity 90.5%; Pred. No. 73;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 acataatgctcgttcacac 21
||||| ||||| ||||| ||||| |||||
Db 546 acatgatgcatcgttcacac 566

RESULT 8
US-09-702-134-41881/C
; Sequence 41881, Application US/09702134
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Plant Genome Sequence and Uses Thereof
; FILE REFERENCE: 38-21(51237)F
; CURRENT APPLICATION NUMBER: US/09702.134
; CURRENT FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 52202
; SEQ ID NO 41881
; LENGTH: 704
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-702-134-41881

Query Match 84.8%; Score 17.8; DB 28; Length 704;
Best Local Similarity 90.5%; Pred. No. 74;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 acataatgctcgttcacac 21
||||| ||||| ||||| ||||| |||||
Db 248 ACATAGGCGCTCGTTCACAC 228

RESULT 9
US-60-173-464-17842
; Sequence 17842, Application US/60173464
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W. D.
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
; FILE REFERENCE: CL000173
; CURRENT APPLICATION NUMBER: US/60/173.464
; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 30269
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17842
; LENGTH: 849
; TYPE: DNA
; ORGANISM: Drosophila
US-60-173-464-17842

Query Match 84.8%; Score 17.8; DB 49; Length 849;
Best Local Similarity 90.5%; Pred. No. 76;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 acataatgctcgttcacac 21
||||| ||||| ||||| ||||| |||||
Db 546 acatgatgcatcgttcacac 566

RESULT 10
US-60-167-217-21922/C
; Sequence 21922, Application US/60167217
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W. D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES
; FILE REFERENCE: CL000152
; CURRENT APPLICATION NUMBER: US/60/167.217
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 23195
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21922
; LENGTH: 3382
; TYPE: DNA
; ORGANISM: Drosophila
US-60-167-217-21922

Query Match 84.8%; Score 17.8; DB 48; Length 3382;
Best Local Similarity 90.5%; Pred. No. 93;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 acataatgctcgttcacac 21
||||| ||||| ||||| ||||| |||||
Db 1109 ACATGATGCGCTCGTTCACAC 1089

RESULT 11
US-60-191-637-21771/C
; Sequence 21771, Application US/60191637
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; FILE REFERENCE: CL000392
; CURRENT APPLICATION NUMBER: US/60/191.637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21771
; LENGTH: 3382
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-60-191-637-21771

Query Match 84.8%; Score 17.8; DB 51; Length 3382;
Best Local Similarity 90.5%; Pred. No. 93;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 acataatgctcgttcacac 21
||||| ||||| ||||| ||||| |||||
Db 1109 ACATGATGCGCTCGTTCACAC 1089

RESULT 12
US-60-191-681-17178/C
; Sequence 17178, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W. D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; FILE REFERENCE: CL000390

```

; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 30973
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17178
; LENGTH: 3382
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-60-191-681-17178

```

```

Query Match      84.8%; Score 17.8; DB 51; Length 3382;
Best Local Similarity 90.5%; Pred. No. 93;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 acataatggcctcgcttcacac 21
    ||||| ||||| ||||| |||||
Db 1109 ACATGATGCGCATGCTTCACAC 1089

```

```

RESULT 13
US-60-173-464-17841/C
; Sequence 17841, Application US/60173464
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W.D.
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C10001173
; CURRENT APPLICATION NUMBER: US/60/173,464
; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 30269
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17841
; LENGTH: 3577
; TYPE: DNA
; ORGANISM: Drosophila
US-60-173-464-17841

```

```

Query Match      84.8%; Score 17.8; DB 49; Length 3577;
Best Local Similarity 90.5%; Pred. No. 94;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 acataatggcctcgcttcacac 21
    ||||| ||||| ||||| |||||
Db 1304 ACATGATGCGCATGCTTCACAC 1284

```

```

RESULT 14
US-09-528-237A-1266
; Sequence 1266, Application US/09528237A
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: The Drosophila Genome, Primary Nucleic
; TITLE OF INVENTION: Acid Sequences, Systems Containing The Nucleic Acid
; TITLE OF INVENTION: Sequences and Uses Thereof
; FILE REFERENCE: C1000284
; CURRENT APPLICATION NUMBER: US/09/528,237A
; CURRENT FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 2926
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1266
; LENGTH: 59140
; TYPE: DNA
; ORGANISM: Drosophila
US-09-528-237A-1266

```

```

Query Match      84.8%; Score 17.8; DB 19; Length 59140;
Best Local Similarity 90.5%; Pred. No. 1,4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 acataatggcctcgcttcacac 21
    ||||| ||||| ||||| |||||
Db 32318 acatgatgcatgcttcacac 32338

```

```

RESULT 15
US-60-212-664-166/C
; Sequence 166, Application US/60212664
; GENERAL INFORMATION:
; APPLICANT: Ladunga, Steve
; APPLICANT: Splier, Gene
; APPLICANT: Greenberg, Simon
; APPLICANT: Rabkin, Steven
; APPLICANT: Wang, Yu
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1000687
; CURRENT APPLICATION NUMBER: US/60/212,664
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 636
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 166
; LENGTH: 183411
; TYPE: DNA
; ORGANISM: HUMAN
; NAME/KEY: misc_feature
; LOCATION: (1)...(183411)
; OTHER INFORMATION: n = A,T,C or G
US-60-212-664-166

```

```

Query Match      82.9%; Score 17.4; DB 53; Length 183411;
Best Local Similarity 94.7%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 2 cataatggcctcgcttcaca 20
    ||||| ||||| ||||| |||||
Db 115270 CATATGCGCTCGTTCAAA 115252

```

```

Search completed: September 21, 2001, 02:20:06
Job time: 30478 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 21, 2001, 02:39:25 : Search time 1035.34 Seconds
(without alignments)
40.480 Million cell updates/sec

Title: US-09-138-735-10

Perfect score: 21

Sequence: 1 acataatggcctcgttcacac 21

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1741792 seqs, 997871973 residues

Total number of hits satisfying chosen parameters: 3483584

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_New:*
1: /cgnl_7/ptodata/2/pna/PCY_NEW_COMB.seq:*
2: /cgnl_7/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgnl_7/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgnl_7/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgnl_7/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgnl_7/ptodata/2/pna/US09_NEW_COMB.seq1:*
7: /cgnl_7/ptodata/2/pna/US09_NEW_COMB.seq2:*
8: /cgnl_7/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|---------|----|---------------------|
| 1 | 21 | 100.0 | 21 | 6 | US-09-138-735-10 |
| 2 | 21 | 100.0 | 3402 | 6 | US-09-138-735-1 |
| 3 | 16.8 | 80.0 | 90824 | 6 | US-09-803-736-570 |
| 4 | 16.4 | 78.1 | 227 | 8 | US-60-253-378-23517 |
| 5 | 16.4 | 78.1 | 285 | 8 | US-60-253-378-22541 |
| 6 | 16.4 | 78.1 | 320 | 8 | US-60-253-456-24803 |
| 7 | 16.4 | 77.1 | 1503841 | 6 | US-09-795-686-1 |
| 8 | 16.2 | 77.1 | 108 | 4 | US-08-798-074B-1985 |
| 9 | 16.2 | 77.1 | 108 | 4 | US-08-798-074B-1985 |
| 10 | 16.2 | 77.1 | 108 | 4 | US-08-798-074C-1985 |
| 11 | 16.2 | 77.1 | 110 | 7 | US-09-764-874-1199 |
| 12 | 16.2 | 77.1 | 149 | 4 | US-08-798-074B-1726 |
| 13 | 16.2 | 77.1 | 149 | 4 | US-08-798-074B-1726 |
| 14 | 16.2 | 77.1 | 149 | 4 | US-08-798-074C-1726 |
| 15 | 15.8 | 75.2 | 326 | 7 | US-09-834-366-48084 |
| 16 | 15.8 | 75.2 | 403 | 8 | US-60-253-652-15179 |
| 17 | 15.8 | 75.2 | 451 | 7 | US-09-834-366-50522 |
| 18 | 15.8 | 75.2 | 99008 | 6 | US-09-803-736-1399 |
| 19 | 15.4 | 73.3 | 421 | 7 | US-09-834-366-11277 |
| 20 | 15.4 | 73.3 | 452 | 6 | US-09-909-629-19092 |
| 21 | 15.4 | 73.3 | 454 | 6 | US-09-909-627-1543 |
| 22 | 15.4 | 73.3 | 794 | 6 | US-09-808-383-5722 |
| 23 | 15.4 | 73.3 | 95173 | 6 | US-09-803-736-1184 |
| 24 | 15.4 | 73.3 | 105680 | 6 | US-09-803-736-545 |
| 25 | 15.4 | 73.3 | 187892 | 6 | US-09-803-110-208 |

| | | | | | | |
|------|------|------|------|---|---------------------|-------------------|
| c 26 | 15.2 | 72.4 | 257 | 8 | US-60-253-456-6471 | Sequence 6471, Ap |
| c 27 | 15.2 | 72.4 | 288 | 6 | US-09-866-555-2028 | Sequence 2028, Ap |
| c 28 | 15.2 | 72.4 | 296 | 7 | US-09-867-716-2330 | Sequence 2330, Ap |
| c 29 | 15.2 | 72.4 | 348 | 7 | US-09-867-716-15052 | Sequence 15052, A |
| c 30 | 15.2 | 72.4 | 376 | 1 | PCT-US01-18569-1889 | Sequence 1889, Ap |
| c 31 | 15.2 | 72.4 | 377 | 8 | US-60-256-867-486 | Sequence 486, Ap |
| c 32 | 15.2 | 72.4 | 402 | 8 | US-60-253-457-43854 | Sequence 43854, A |
| c 33 | 15.2 | 72.4 | 410 | 6 | US-09-790-435-1341 | Sequence 1341, Ap |
| c 34 | 15.2 | 72.4 | 425 | 6 | US-09-909-637-14034 | Sequence 14034, A |
| c 35 | 15.2 | 72.4 | 470 | 6 | US-09-796-632-5529 | Sequence 5529, Ap |
| c 36 | 15.2 | 72.4 | 508 | 8 | US-60-253-652-24830 | Sequence 24830, A |
| c 37 | 15.2 | 72.4 | 598 | 8 | US-60-253-651-22731 | Sequence 22731, A |
| c 38 | 15.2 | 72.4 | 637 | 7 | US-09-910-689-189 | Sequence 189, Ap |
| c 39 | 15.2 | 72.4 | 707 | 7 | US-09-867-716-19160 | Sequence 19160, A |
| c 40 | 15.2 | 72.4 | 717 | 7 | US-09-824-518-312 | Sequence 312, Ap |
| c 41 | 15.2 | 72.4 | 1046 | 8 | US-60-312-544-4380 | Sequence 4380, Ap |
| c 42 | 15.2 | 72.4 | 1968 | 6 | US-09-764-891-7620 | Sequence 7620, Ap |
| c 43 | 15.2 | 72.4 | 3417 | 6 | US-09-879-238-1 | Sequence 1, Appl |
| c 44 | 15.2 | 72.4 | 3456 | 6 | US-09-764-856-42 | Sequence 42, Appl |
| c 45 | 15.2 | 72.4 | 4208 | 8 | US-60-278-258-4178 | Sequence 4178, Ap |

ALIGNMENTS

RESULT 1
US-09-138-735-10
Sequence 10, Application US/09138735
GENERAL INFORMATION:
APPLICANT: PARANHOS-BACCALA, Glaucia
APPLICANT: LESNECHAL, Mylene
APPLICANT: JOLIVER, Michel
TITLE OF INVENTION: TRYPANOSOMA CRUZI ANTIGEN, GENE ENCODING THEREFOR AND METHODS
FILE REFERENCE: WPB 36400B
CURRENT APPLICATION NUMBER: US/09/138,735
PRIOR FILING DATE: 1998-08-24
PRIOR APPLICATION NUMBER: US 08/480,917
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: FR 94/10132
PRIOR FILING DATE: 1994-08-12
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn version 3.0
SEQ ID NO 10
LENGTH: 21
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: probe/primer
US-09-138-735-10
Query Match 100.0%; Score 21; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.099;
Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 acataatggcctcgttcacac 21
1 acataatggcctcgttcacac 21
RESULT 2
US-09-138-735-1/c
Sequence 1, Application US/09138735
GENERAL INFORMATION:
APPLICANT: PARANHOS-BACCALA, Glaucia
APPLICANT: LESNECHAL, Mylene
APPLICANT: JOLIVER, Michel
TITLE OF INVENTION: TRYPANOSOMA CRUZI ANTIGEN, GENE ENCODING THEREFOR AND METHODS
FILE REFERENCE: WPB 36400B
CURRENT APPLICATION NUMBER: US/09/138,735
CURRENT FILING DATE: 1998-08-24

;; PRIOR APPLICATION NUMBER: US 08/480,917
;; PRIOR FILING DATE: 1995-06-07
;; PRIOR APPLICATION NUMBER: FR 94/10132
;; PRIOR FILING DATE: 1994-08-12
;; NUMBER OF SEQ ID NOS: 13
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 1
;; LENGTH: 3402
;; TYPE: DNA
;; ORGANISM: Trypanosoma cruzi
US-09-138-735-1

Query Match 100.0%; Score 21; DB 6; Length 3402;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 acataatgacctgttcacac 21
DB 2207 ACATTAATGGCCTGTCACAC 2187

RESULT 3
US-09-803-736-570/c
;; Sequence 570, Application US/09803736
;; GENERAL INFORMATION:
;; APPLICANT: Bush, David F.
;; APPLICANT: Levin, Irena M.
;; APPLICANT: Norris, Susan R.
;; APPLICANT: Rounsley, Steven D.
;; APPLICANT: Wiegand, Roger C.
;; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
;; FILE REFERENCE: 38-10(15493)D
;; CURRENT APPLICATION NUMBER: US/09/803,736
;; CURRENT FILING DATE: 2001-03-12
;; PRIOR APPLICATION NUMBER: US 09/534,859
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: Identified by Attorney Docket number 04983.0206CPUS01 38-10
;; PRIOR FILING DATE: 2000-10-20
;; NUMBER OF SEQ ID NOS: 1582
;; SEQ ID NO 570
;; LENGTH: 90824
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
US-09-803-736-570

Query Match 80.0%; Score 16.8; DB 6; Length 90824;
Best Local Similarity 90.0%; Pred. No. 75;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 cataatgacctgttcacac 21
DB 43650 CATTAATGGCCTGCTTAAC 43631

RESULT 4
US-60-253-378-23517/c
;; Sequence 23517, Application US/60253378
;; GENERAL INFORMATION:
;; APPLICANT: Havukkala, Ilkka J
;; TITLE OF INVENTION: Polynucleotides isolated from plants
;; FILE REFERENCE: 1054P3
;; CURRENT APPLICATION NUMBER: US/60/253,378
;; CURRENT FILING DATE: 2000-11-27
;; NUMBER OF SEQ ID NOS: 40367
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 23517
;; LENGTH: 227
;; TYPE: DNA
;; ORGANISM: Pinus radiata
US-60-253-378-23517

Query Match 78.1%; Score 16.4; DB 8; Length 227;
Best Local Similarity 94.4%; Pred. No. 44;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 taatgacctgttcacac 21
DB 102 TAATGGCCTGTCACAC 85

RESULT 5
US-60-253-378-22541/c
;; Sequence 22541, Application US/60253378
;; GENERAL INFORMATION:
;; APPLICANT: Havukkala, Ilkka J
;; TITLE OF INVENTION: Polynucleotides isolated from plants
;; FILE REFERENCE: 1054P3
;; CURRENT APPLICATION NUMBER: US/60/253,378
;; CURRENT FILING DATE: 2000-11-27
;; NUMBER OF SEQ ID NOS: 40367
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 22541
;; LENGTH: 285
;; TYPE: DNA
;; ORGANISM: Pinus radiata
US-60-253-378-22541

Query Match 78.1%; Score 16.4; DB 8; Length 285;
Best Local Similarity 94.4%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 taatgacctgttcacac 21
DB 102 TAATGGCCTGTCACAC 85

RESULT 6
US-60-253-456-24803/c
;; Sequence 24803, Application US/60253456
;; GENERAL INFORMATION:
;; APPLICANT: Havukkala, Ilkka J
;; TITLE OF INVENTION: Polynucleotides, isolated from
;; FILE REFERENCE: 1054P1
;; CURRENT APPLICATION NUMBER: US/60/253,456
;; CURRENT FILING DATE: 2000-11-27
;; NUMBER OF SEQ ID NOS: 37096
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 24803
;; LENGTH: 320
;; TYPE: DNA
;; ORGANISM: Pinus radiata
US-60-253-456-24803

Query Match 78.1%; Score 16.4; DB 8; Length 320;
Best Local Similarity 94.4%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 taatgacctgttcacac 21
DB 19 TAATGGCCTGTCACAC 2

RESULT 7
US-09-795-686-1
;; Sequence 1, Application US/09795686
;; GENERAL INFORMATION:
;; APPLICANT: Stefansson, Hreinn
;; APPLICANT: Steinhorsdottir, Valgerdur


```
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345, 2005-001
; CURRENT APPLICATION NUMBER: US/09/795,686
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,715
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1503841
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: r=g or a
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: y=t/u or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: m=a or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: k=g or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: s=g or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: w=a or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: d=a or g or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: n=a or c or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: v=a or g or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: n=a or g or c or t/u
; US-09-795-686-1

Query Match          78.1%; Score 16.4; DB 6; Length 1503841;
Best Local Similarity 94.4%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 acataatgacctcgttca 18
    ||| ||||| ||||| |||||
Db 243742 acaaatgacctcgttca 243759
```

```
RESULT 8
US-08-798-074B-1985
; Sequence 1985, Application US/08798074B
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Li, Haodong
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 22
; FILE REFERENCE: PO-22
; CURRENT APPLICATION NUMBER: US/08/798,074B
; CURRENT FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: 60/011,617
; PRIOR FILING DATE: 1996-02-13
```

```
; NUMBER OF SEQ ID NOS: 13449
; SEQ ID NO 1985
; LENGTH: 108
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (31)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (76)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (80)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (94)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (107)
; OTHER INFORMATION: n equals a,t,g, or c
; US-08-798-074B-1985
```

```
Query Match          77.1%; Score 16.2; DB 4; Length 108;
Best Local Similarity 85.7%; Pred. No. 50;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 acataatgacctcgttcacac 21
    |||| ||||| ||||| |||||
Db 42 acataatgacctcgttcacgc 62
```

```
RESULT 9
US-08-798-074B-1985
; Sequence 1985, Application US/08798074B
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Li, Haodong
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 22
; FILE REFERENCE: PO-22
; CURRENT APPLICATION NUMBER: US/08/798,074B
; CURRENT FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: 60/011,617
; PRIOR FILING DATE: 1996-02-13
; NUMBER OF SEQ ID NOS: 13449
; SEQ ID NO 1985
; LENGTH: 108
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (31)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (76)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (80)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (94)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (107)
```

OTHER INFORMATION: n equals a,t,g, or c
US-08-798-074B-1985

Query Match 77.1%; Score 16.2; DB 4; Length 108;
Best Local Similarity 85.7%; Pred. No. 30;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 acataatgacctgttcacac 21
|||||
Db 42 acattatgacctgttaacgc 62

RESULT 10

US-08-798-074C-1985
Sequence 1985, Application US/08798074C
GENERAL INFORMATION:
APPLICANT: Rosen, Craig A.
APPLICANT: Dillon, Patrick J.
APPLICANT: Li, Haodong
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 22
FILE REFERENCE: PO-22
CURRENT APPLICATION NUMBER: US/08/798,074C
CURRENT FILING DATE: 1997-02-12
PRIOR APPLICATION NUMBER: 60/011,617
PRIOR FILING DATE: 1996-02-13
NUMBER OF SEQ ID NOS: 13449
SEQ ID NO 1985

LENGTH: 108
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (8)
OTHER INFORMATION: n equals a,t,g, or c
LOCATION: (31)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (76)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (80)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (94)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (107)
OTHER INFORMATION: n equals a,t,g, or c
US-08-798-074C-1985

Query Match 77.1%; Score 16.2; DB 4; Length 108;
Best Local Similarity 85.7%; Pred. No. 30;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 acataatgacctgttcacac 21
|||||
Db 42 acattatgacctgttaacgc 62

RESULT 11

US-09-764-874-1199
Sequence 1199, Application US/09764874
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC010
CURRENT APPLICATION NUMBER: US/09/764,874
CURRENT FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065

PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/225,757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01

;; PRIOR APPLICATION NUMBER: 60/229,287
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 60/229,513
;; PRIOR FILING DATE: 2000-09-05
;; PRIOR APPLICATION NUMBER: 60/231,413
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/229,509
;; PRIOR FILING DATE: 2000-09-05
;; PRIOR APPLICATION NUMBER: 60/236,367
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/237,039
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,038
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/236,370
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/236,802
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,037
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,040
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/240,960
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/239,935
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: 60/239,937
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: 60/241,787
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/246,474
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/246,532
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/249,216
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,210
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/226,681
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/225,759
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/225,213
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/227,182
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/225,214
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/235,836
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/230,438
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/215,135
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: 60/225,266
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/249,218
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,208
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,213
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,212
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,207
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,245
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,244
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,217
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,211

;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,215
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,264
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,214
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,297
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/232,400
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/231,242
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,081
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,080
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,414
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,244
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/233,064
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/233,063
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,397
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,399
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,401
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/241,808
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,826
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,786
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,221
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/246,475
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/231,243
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/233,065
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,398

Query Match 77.1%; Score 16.2; DB 7; Length 110;
Best Local-Similarity 85.7%; Pred. No. 50;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 acataatgctcgttacac 21
||||||| 11 111 1111
Db 73 acataatgcccgttacac 93

RESULT 12
US-08-798-074B-1726
; Sequence 1726, Application US/08798074B
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Li, Haodong
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 22
; FILE REFERENCE: PO-22
; CURRENT APPLICATION NUMBER: US/08/798, 074B
; PRIOR FILING DATE: 1997-02-12
; PRIOR APPLICATION NUMBER: 60/011,617
; PRIOR FILING DATE: 1996-02-13
; NUMBER OF SEQ ID NOS: 13449
; SEQ ID NO 1726
; LENGTH: 149

```
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (11)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (16)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (19)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (20)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (25)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (109)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (112)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (113)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (126)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (127)
OTHER INFORMATION: n equals a,t,g, or c
US-08-798-074B-1726
```

Query Match 77.1%; Score 16.2; DB 4; Length 149;
Best Local Similarity 85.7%; Pred. No. 52;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 acataatggcctcgtcacac 21
    ||||| ||||| ||||| |||
DB 78 acatatggcctcgttaacgc 98
```

```
RESULT 13
US-08-798-074B-1726
; Sequence 1726, Application US/08798074B
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Li, Haodong
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 22
; FILE REFERENCE: PO-22
; CURRENT APPLICATION NUMBER: US/08/798,074B
; PRIOR APPLICATION NUMBER: 60/011,617
; PRIOR FILING DATE: 1996-02-13
; NUMBER OF SEQ ID NOS: 13449
; SEQ ID NO 1726
; LENGTH: 149
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (16)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (19)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (20)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (25)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (109)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (112)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (113)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (126)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (127)
; OTHER INFORMATION: n equals a,t,g, or c
US-08-798-074B-1726
```

```
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (20)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (25)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (109)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (112)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (113)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (126)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (127)
OTHER INFORMATION: n equals a,t,g, or c
US-08-798-074B-1726
```

Query Match 77.1%; Score 16.2; DB 4; Length 149;
Best Local Similarity 85.7%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 acataatggcctcgtcacac 21
    ||||| ||||| ||||| |||
DB 78 acatatggcctcgttaacgc 98
```

```
RESULT 14
US-08-798-074C-1726
; Sequence 1726, Application US/08798074C
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Li, Haodong
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 22
; FILE REFERENCE: PO-22
; CURRENT APPLICATION NUMBER: US/08/798,074C
; PRIOR APPLICATION NUMBER: 60/011,617
; PRIOR FILING DATE: 1996-02-13
; NUMBER OF SEQ ID NOS: 13449
; SEQ ID NO 1726
; LENGTH: 149
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (16)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (19)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (20)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (25)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (109)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (112)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (113)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (126)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (127)
; OTHER INFORMATION: n equals a,t,g, or c
US-08-798-074C-1726
```

```
LOCATION: (112)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc.feature
LOCATION: (113)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc.feature
LOCATION: (126)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc.feature
LOCATION: (127)
OTHER INFORMATION: n equals a,t,g, or c
US-08-798-074C-1726
```

```
Query Match 77.1%; Score 16.2; DB 4; Length 149;
Best Local Similarity 85.7%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 1 acataatggcctcgltcacac 21
    ||| ||||| ||| |||
Db 78 acataatggcctcgltacgc 98
```

```
RESULT 15
US-09-834-366-48084
Sequence 48084, Application US/09834366
GENERAL INFORMATION:
APPLICANT: Benjamin, Stephane
APPLICANT: Tanaka, Hiroaki
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Jobert, Severin
APPLICANT: Giordano, Jean-Yves
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: 81.US2.REG
CURRENT APPLICATION NUMBER: US/09/834,366
CURRENT FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/197,873
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 52153
SOFTWARE: patent.pm
SEQ ID NO 48084
LENGTH: 326
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 26..27,214
OTHER INFORMATION: n-a, g, c or t
US-09-834-366-48084
```

```
Query Match 75.2%; Score 15.8; DB 7; Length 326;
Best Local Similarity 81.0%; Pred. No. 98;
Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 1 acataatggcctcgltcacac 21
    ||||| :||| |||||
Db 275 acataaascctcatcacac 295
```

Search completed: September 21, 2001, 02:40:43
Job time: 15129 sec

THIS PAGE BLANK (USPTO)